

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
імені ІГОРЯ СІКОРСЬКОГО»**

Інститут телекомунікаційних систем

Кафедра Інформаційно-телекомунікаційних мереж

До захисту допущено:

Завідувач кафедри

_____ Лариса ГЛОБА

«__» _____ 2020 р.

Дипломна робота

на здобуття ступеня бакалавра

за освітньо-професійною програмою «Інформаційно-комунікаційні технології»

спеціальності 172 «Телекомунікації та радіотехніка»

**на тему: «Інтелектуальні методи обробки різномірної слабоструктурованої
інформації»**

Виконав:

студент IV курсу, групи ТІ-61

Задоєнко Богдан Олександрович _____

Керівник:

професор кафедри ІТМ, д.т.н., професор,

Глоба Лариса Сергіївна _____

Консультант:

доцент кафедри ІТМ, к.т.н., доцент,

Новогрудська Ріна Леонідівна _____

Рецензент:

доцент кафедри ТК, к.т.н., доцент,

Міночкін Дмитро Анатолійович _____

Засвідчую, що у цій дипломній роботі
немає запозичень з праць інших авторів
без відповідних посилань.

Студент _____

Київ – 2020 року

Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Інститут телекомунікаційних систем
Кафедра Інформаційно-телекомунікаційних мереж

Рівень вищої освіти – перший (бакалаврський)

Спеціальність – 172 «Телекомунікації та радіотехніка»

Освітньо-професійна програма «Інформаційно-комунікаційні технології»

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Лариса ГЛОБА

«___» _____ 2020 р.

ЗАВДАННЯ

на дипломну роботу студенту

Задосенку Богдану Олександровичку

1. Тема роботи «Інтелектуальні методи обробки різномірної слабоструктурованої інформації», керівник роботи професор кафедри інформаційно-телекомунікаційних мереж Глоба Лариса Сергіївна, д.т.н., професор, затверджені наказом по університету від «30» березня 2020 р. № 924-с
2. Термін подання студентом роботи 8 червня 2020 р.
3. Вихідні дані до роботи:
 1. Методика оцінювання ефективності діяльності наукових установ.
 2. Моделі та методи обробки різномірної слабоструктурованої інформації.
4. Зміст роботи:
 1. Провести порівняльний аналіз моделей представлення інформації.
 2. Провести огляд процесу онтологічного моделювання інформації.
 3. Проаналізувати показники діяльності наукових та освітніх установ, та критерії оцінки ефективності їх діяльності.

4. Запропонувати модель представлення різномірної слабоструктурованої інформації наукових та освітніх установ.
 5. Описати елементи моделі представлення різномірної слабоструктурованої інформації.
 6. Розробити онтологічну модель для представлення різномірної слабоструктурованої інформації наукових та освітніх установ засобами платформи ТОДАОС.
5. Перелік ілюстративного матеріалу (із зазначенням плакатів, презентацій тощо)
1. Тема, актуальність, мета, задачі.
 2. Огляд інтелектуальних моделей представлення інформації.
 3. Порівняльний аналіз моделей представлення інформації.
 4. Аналіз показників діяльності наукових та освітніх установ, та критерії оцінки ефективності їх діяльності.
 5. Опис елементів моделі представлення різномірної слабоструктурованої інформації.
 6. Опис розробленої онтологічної моделі для представлення різномірної слабоструктурованої інформації наукових та освітніх установ засобами платформи ТОДАОС.
 7. Загальні висновки.

6. Консультанти розділів роботи*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Представлення різномірної слабоструктурованої інформації наукових установ	Новогрудська Р.Л., доцент кафедри ІТМ		

7. Дата видачі завдання: 7 жовтня 2019 р.

Календарний план

№ з/п	Назва етапів виконання дипломної роботи	Термін виконання етапів роботи	Примітка
1	Дослідження та вивчення отриманого завдання	07.10.19 – 10.11.19	
2	Вивчення основних задач, корті можуть бути вирішені шляхом інтелектуальної обробки інформації.	11.11.19 – 25.01.19	
3	Аналіз існуючих методів та моделей інтелектуальної обробки різномірної слабоструктурованої інформації.	26.01.2020 – 01.03.2020	
4	Попередній збір показників діяльності наукових установ.	02.03.2020 – 20.03.2020	
5	Аналіз критеріїв оцінки ефективності діяльності наукових установ.	21.03.2020 – 05.04.2020	
6	Розробка елементів моделі представлення різномірної слабоструктурованої інформації.	06.04.2020 – 25.04.2020	
7	Розробка онтологічної моделі для представлення різномірної слабоструктурованої інформації наукових та освітніх установ засобами платформи ТОДАОС.	27.04.2020 – 10.05.2020	
8	Підготовка тексту диплому	27.04.2020 – 06.06.2020	

Студент

Богдан ЗАДОЄНКО

Керівник

Лариса ГЛОБА

РЕФЕРАТ

Робота містить 66 сторінок, 20 рисунків, 1 таблицю. Було використано 27 джерел.

Мета роботи: підвищення якості оцінки, функціонування освітніх установ, за рахунок систематизації та структуризації різномірної слабоструктурованої інформації.

В ході виконання даної роботи проведено огляд моделей інтелектуальної обробки різномірної слабоструктурованої інформації. Проведено аналіз моделей обробки інформації. Проведено вибірку показників та критеріїв оцінювання наукових установ. Запропоновано модель інтелектуальної обробки інформації. Описано елементи моделі представлення різномірної слабоструктурованої інформації. Розроблено та описано онтологічну модель представлення інформації наукових установ.

Ключові слова: Інтелектуальні методи обробки інформації, слабоструктурована інформація, оцінювання наукових установ, онтологічна модель, представлення інформації.

ABSTRACT

The work contains 66 pages, 20 figures, 1 table. 27 sources were used.

Goal: to improve the quality of assessment, the functioning of educational institutions, through the systematization and structuring of heterogeneous poorly structured information.

In the course of this work, a review of models of intelligent processing of heterogeneous poorly structured information. The analysis of information processing models is carried out. A sample of indicators and evaluation criteria of scientific institutions was conducted. A model of intelligent information processing is proposed. The elements of the model of representation of heterogeneous poorly structured information are described. An ontological model of presenting information of scientific institutions is developed and described.

Keywords: Intelligent methods of information processing, poorly structured information, evaluation of scientific institutions, ontological model, information presentation.

ЗМІСТ

ВСТУП.....	8
РОЗДІЛ 1.....	10
ПІДХОДИ ДО ПРЕДСТАВЛЕННЯ РІЗНОРІДНОЇ СЛАБОСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ	10
1.1. Огляд моделей представлення інформації	10
1.2. Порівняльний аналіз моделей представлення інформації.....	26
Висновки	32
РОЗДІЛ 2.....	33
ПРЕДСТАВЛЕННЯ РІЗНОРІДНОЇ СЛАБОСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ НАУКОВИХ УСТАНОВ	33
2.1. Характерні особливості інформації наукових установ	33
2.2. Модель представлення інформації наукових установ	35
2.3. Елементи системи онтологій для представлення інформації наукових установ	38
Висновки	42
РОЗДІЛ 3.....	43
ОНТОЛОГІЯ РІЗНОРІДНОЇ СЛАБОСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ НАУКОВИХ УСТАНОВ	43
3.1. Програмні середовища проектування онтологічних моделей	43
3.2. Характерні особливості програмної платформи ТОДАОС	45
3.3. Онтологія інформації наукових установ на базі платформи ТОДАОС	57
Висновки	62
ЗАГАЛЬНІ ВИСНОВКИ ПО РОБОТІ.....	63
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	64

ВСТУП

Актуальність: Наукова діяльність є одним із найбільш впливових чинників розвитку економіки кожної країни. Об'єктивний аналіз ефективності діяльності наукових установ та значення їх діяльності з точки зору світової та вітчизняної науки, освіти, економіки, культури, підвищення якості життя людини є вкрай актуальним. Важливим є отримання та аналіз даних відкритим та прозорим способом із застосуванням об'єктивних методів отримання показників діяльності.

Державні установи такі як міністерства та відомства відносяться до складних організаційних систем. Вони мають розвинену ієрархічну структуру, що складається з багатьох підрозділів та підлеглих установ. Інформаційні технології складають основу інформаційно-аналітичної діяльності для задач оцінювання таких об'єктів, а обробка інформації є однією з найбільш складних частин процесу оцінювання. Це обумовлено великою кількістю складних інформаційних потоків, що циркулюють в таких об'єктах, високою динамікою зміни їх структури, яка обумовлена законодавчими змінами та наявністю на кожному рівні суб'єктивних факторів, необхідністю врахування при обробці як кількісної так і якісної інформації.

Обробка інформації в таких установах виконується експертами, які, базуючись на власному досвіді та за допомогою спрощених математичних методів, проводять оцінку загального стану результатів діяльності. При цьому необхідно визначити ключові моменти такої діяльності :

- оцінювана інформація часто є неповною, частина даних подається в кількісному вигляді, частина – в якісному;
- змінність вимог до проведення обробки інформації і до моделі системи;
- робота з оцінки ситуації виконується людиною на всіх рівнях ієрархії об'єкту, тому навіть в однотипних підрозділах при однакових вхідних даних можуть бути отримані різні оцінки;

- внаслідок того, що обробка інформації здійснюється людиною така система втрачає оперативність;
- виходячи з кінцевих звітів, які подаються керівнику установи, важко простежити та визначити ті параметри, які найбільше впливають на такий стан справ.

Внаслідок цього є актуальною наукова-технічна задача розробки підходу до обробки інформації в складних організаційних системах, який би включав метод обробки інформації на основі деревоподібних баз знань, який використовує змішану, а саме чітку і нечітку, схему виведення [1].

Об'єкт дослідження – процес оцінювання показників діяльності наукових установ.

Предмет дослідження – моделі та методи представлення та обробки різномірної слабоструктурованої інформації .

Мета і задачі дослідження: Підвищення якості оцінки, функціонування освітніх установ, за рахунок систематизації та структуризації різномірної слабоструктурованої інформації.

Для досягнення мети дослідження було поставлено та вирішено такі основні задачі:

1. Провести порівняльний аналіз моделей представлення інформації.
2. Провести огляд процесу онтологічного моделювання інформації.
3. Проаналізувати показники діяльності наукових та освітніх установ, та критерії оцінки ефективності їх діяльності.
4. Запропонувати модель представлення різномірної слабоструктурованої інформації наукових та освітніх установ.
5. Описати елементи моделі представлення різномірної слабоструктурованої інформації.
6. Розробити онтологічну модель для представлення різномірної слабоструктурованої інформації наукових та освітніх установ засобами платформи ТОДАОС.

РОЗДІЛ 1.

ПІДХОДИ ДО ПРЕДСТАВЛЕННЯ РІЗНОРІДНОЇ СЛАБОСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ

Представлення знань - це один з важливих напрямків досліджень в області інтелектуальної обробки різноманітної слабоструктурованої інформації. При їх створенні не можна обійтися без різноманітності моделей представлення знань. Це пояснюється, тим фактом, що без знань обробка інформації не може існувати в принципі.

Не дивлячись на той факт, що на сьогоднішній день існує велика кількість способів представлення знань, питання вибору конкретного методу для вирішення різноманітних задач, залишається актуальним. Існує проблема в поданні знань - як обробляти і зберігати знання в інформаційних системах так, щоб можна було їх використовувати для досягнення поставленої задачі.

1.1. Огляд моделей представлення інформації

Моделі представлення знань забезпечують представлення різноманітних даних, знань та інформації у специфічній формі, прийнятній для подальшої обробки.

Існують наступні моделі представлення знань:

- Фреймова модель
- Мережна модель
- Продукційна модель
- Логічна (предикатна) модель
- Онтологічна модель

Наведемо опис кожної з моделей.

Фреймова модель. . Фреймова модель ґрунтується на такому понятті як фрейм (англ. frame – рамка, каркас). Фреймове представлення є важливим формалізмом представлення знань що дозволяє нам показати поняття

спадкування. Техніка фрейму включає в себе: кількість кадрів або вузлів, які пов'язані між собою відносинами. Кожен фрейм пояснює і екземпляр, і кадр класу. Ідея фрейму спочатку була представлений Марвіном Мінським у 1975 р. як основний спосіб показу діапазону знання.

Фрейм - це група властивостей, що ідентифікують стан об'єкта, і це об'єкт пов'язаний з іншими кадрами або об'єктами. Насправді фрейм - це більше, ніж лише запис або, можливо, структура даних, що містить дані.

У штучному інтелекті фрейм відомий як метод подання знань із заповненням слотів.

На сьогоднішній день відомо, що екземпляр, як правило, є "об'єктом". У цьому випадку об'єкт може бути фізичним об'єктом, однак він не став. Об'єктом може бути власність (наприклад, форма чи колір) або місцеположення, сценарій чи емоція.

Популярні атрибути в представленні знань використовуються на основі фреймового методу:

- 1) Структура фреймів ніби складається з фреймів, розташованих в ієрархічній формі;
- 2) Основні складові фрейми - це слоти, і наповнювачі, які використовуються для цих слотів, повинні бути вказані;
- 3) Характеристики (наповнювачі, обмеження наповнювача тощо), як правило, успадковуються супер-фрейми, щоб мати можливість використовувати "підфрейми" зі структури відповідно до кількох прийомів успадкування.

Фрейм можна описати набором слотів. Кожен слот пояснює певну особливість або процедуру з фрейму. Слоти використовуються для збереження значень. Слот, може включати значення за замовчуванням для різних фреймів і набір принципів, за допомогою яких можна отримати фактичні значення слота.

Слоти можуть бути термінальними або бути фреймами, таким чином утворюючи цілу ієрархічну мережу. Фрейм має майже однорідну структуру і складається зі стандартних одиниць. Кожна така одиниця (слот) містить назву і має своє значення [2].

Формально фрейм – це тип даних виду:

$$F = \langle N, S_1, S_2, S_3 \rangle,$$

де N – ім'я об'єкту;

S_1 – множина слотів, що містять факти, які визначають декларативну семантику фрейму;

S_2 – множина слотів, що забезпечують зв'язки з іншими фреймами (каузальні, семантичні і т. д.);

S_3 – множина слотів, що забезпечують перетворення, які визначають процедурну семантику кадру.

Склад фреймів і слотів в кожній конкретній фреймовій моделі може бути різний, проте в рамках однієї системи доцільно єдине представлення для усунення зайвого ускладнення [3]. Глибина вкладеності слотів у фреймі (число рівнів) залежить від предметної області та мови, що реалізує модель.

Сукупність упорядкованих фреймів називається сценарієм. Структура фрейму зазвичай використовується для «розуміння» природних мов, а також для опису механічних пристроїв та процесу їх функціонування. Це пов'язано з тим, що механічні пристрої складаються з великої кількості елементів, які пов'язані між собою. Кожен фрейм може зберігати інформацію про конкретний елемент, а сукупність фреймів поєднаних між собою буде сценарієм, представляючим певний механічний пристрій [4].

У слот фрейму включені наступні загальновідомі знання:

- 1) Назва кадру;
- 2) З'єднання одного кадру з іншими кадрами.
- 3) Значення слотів: значення слотів може бути булевим, числовим або символічним. Значення слота зазвичай розподіляється під час створення кадру або в рамках процедури під час використання експертних систем;

4) Невиконання значень слотів: це фактично правильно, хоча жодних доказів протилежного не виявлено;

5) Діапазон значень слота: Поле значення слота визначає, чи відповідає конкретний об'єкт стереотипним потребам, окресленим рамкою.

6) Процедурні знання: у слоті є пов'язана з ним процедура, і це здійснюється, коли значення слота потрібно або змінено;

7) Система на основі кадру забезпечує розширення для побудови значення слота за допомогою граней. Фасет - це дійсно спосіб надання розширених знань, що стосується атрибута фрейму. Фасети можна використовувати для встановлення значення атрибута, для управління запитами кінцевих користувачів тощо.

Процедури, що додаються до фреймів, дають нам гнучку організовану основу для обчислень. Обґрунтування в рамках системи фрейму зазвичай починається з того, що система "розпізнає" об'єкт як екземпляр загального фрейму, а потім застосовує процедури, ініційовані цим розпізнаванням. Такі виклики процедур можуть потім отримати більше даних або зміни в базі знань, які можуть відноситися до інших викликів процедур. Коли більше процедур не застосовується, система зупиняється.

Більш конкретно, базовий цикл міркування в рамковій системі має ці три етапи:

1. Користувач або зовнішня система, що використовує систему кадру, оскільки її представлення знань заявляє, що існує об'єкт або ситуація, тим самим створюючи певний загальний кадр;

2. Успадковуються будь-які заповнювачі слотів, які не надані явно, але можуть бути успадковані новим екземпляром кадру;

3. Для кожного слота з наповнювачем запускається будь-яка процедура IF-ADDED, яка може бути успадкована, можливо, спричиняючи заповнення нових слотів або створення нових кадрів, і цикл повторюється

Якщо користувач зовнішня система або додана процедура вимагає заповнення слота, то ми отримуємо таку поведінку:

1. Якщо в слоті зберігається наповнювач, то це значення повертається;
2. В іншому випадку будь-яка процедура IF-NEEDED, яка може бути успадкована, запускається, обчислюючи наповнювач для слота, але, можливо, також спричиняючи інші слоти, які потрібно заповнити, або нові фрейми, які потрібно створити.

Якщо жоден з цих результатів не дає результату, то значення слота вважається невідомим. Зауважимо, що в цьому рахунку спадкування о

Значення властивостей проводиться в момент створення окремого кадру, але процедури IF-NEEDED, які обчислюють значення властивостей, викликаються лише по мірі необхідності. Можливі й інші схеми.

Сюди входить локальне міркування, що включає один кадр. Створюючи фреймову базу знань, можна також подумати про те глобальна структура та обчислення повинні виробляти бажані загальні міркування. Як правило, загальні кадри створюються для будь-якого основного типу об'єкта або типу ситуації, який необхідний для вирішення проблеми. Будь-які обмеження між слотами виражаються доданими процедурами IF-ADDED та IF-NEEDED.

У цьому випадку значення за замовчуванням заповнюються, коли вони доступні на слотах. Варто зазначити, що в оригінальному, психологічному погляді, що спочатку спричинило фрейми, за замовчуванням вважалося, що вони відіграють головну роль у розпізнаванні сцени, ситуації та об'єкта; Вважалося, що люди схильні узагальнювати ситуацію, яку вони бачили раніше, і вважають, що об'єкти та ситуації є "типовими" - мали ключові аспекти, приймаючи їх звичайні значення за замовчуванням - якщо тільки конкретні особливості в окремий випадок був помічений як винятковий.

В цілому, враховуючи обмеження між слотами, які виконуються за допомогою доданих процедур, ми можемо вважати фреймову базу знань як символічну «електронну таблицю» з обмеженнями між об'єктами, про які ми піклуємося, щоб їх розповсюджували доданими процедурами. Але процедури в фреймі можуть зробити набагато більше, включаючи виклик складних дій системою.

Мережна модель представлення знань. Семантичні мережі - це структуровані уявлення знань, які використовуються для міркування та умовиводу. З різних галузей академічних наук та промисловості з'явилося велике різноманіття теорій, моделей, методів та практичних застосувань для створення та використання смислових мереж. Ще в 1975 році Вудс заявив, що семантичні мережі є привабливим поняттям, але їм не вистачає теоретичного обґрунтування та суворості в репрезентативних конвенціях. Відтоді технічний прогрес, особливо поява Інтернету, призвів до експоненціального збільшення доступності даних, які необхідно перетворити на знання та належним чином встигнути забезпечити універсальне використання. Сьогодні багато підходів до семантичних мереж все ще мають мало спільного, крім того, щоб назвати об'єкт дослідження семантичною мережею.

Семантичні мережі потребують трьох складових частин:

- Синтаксис, який визначає типи вузлів та ребер, які можна врахувати.
- Вказання значення або семантики, яку можуть представляти вузли, посилання та вся мережа.
- Правила умовиводу.

Структури даних, що представляють семантичні мережі, містять, як мінімум, вузли, які називаються поняттями, і ребра між поняттями. Поняття - це абстрактні уявлення про ідеї, думки та одиниці знань і сенсу, які люди осмислюють у своїй свідомості. Поняття також можуть бути референтами об'єктів. Прикладами понять є "аналіз соціальних мереж", "співпраця" та "спільність практики". Коли поняття мають представлення природною мовою, відповідне слово або послідовність слів використовують як мітки для вузлів у семантичних мережах. Якщо посилання в семантичних мережах набираються, тип зв'язку пояснює характер взаємозв'язку між підключеними вузлами. В іншому випадку посилання представляють або встановлюють якусь змістовну асоціацію між парами понять. Формування посилань у семантичних мережах сильно залежить від типу вхідних даних та

передбачуваного використання мережі, як описано далі в цьому записі. Крім цих обмежень, уявлення семантичних мереж можуть знаходитися в будь-якому місці між строго формалізованим та неформальним.

У багатьох різних підходах найменша структурована одиниця семантичної мережі, як правило, є трійкою, яка описує джерело чи предмет, предикат та об'єкт чи ціль дії. Ця основна структура використовується, наприклад, для опису того, хто робив, кому для представлення даних про події, і хто сказав, що кому для комунікаційних даних. Такі трійки можуть бути додатково покращені фоновою інформацією, такою як просторові та часові дані, атрибутами вузлів та посилань.

Семантичні мережі є результатом перекладу або перетворення даних, з яких були побудовані мережі. Переклади перетворюють природне введення мови в ізоморфні, структуровані уявлення, які використовуються як вхід до механізмів висновку. Цей підхід розглядався в Штучному інтелекті як одна спроба зрозуміти природну мову, і він показав, що він може бути використаний для невеликих вкладних даних. Трансформації - це процеси абстрагування, які використовуються для збереження та розкриття сутностей та відносин, які явно або неявно представлені у вхідних даних. Мета з перетвореннями - зменшення розмірності вхідних даних для того, щоб зафіксувати відповідні структурні взаємозв'язки та зробити мережеві дані доступними для висновку.

Семантичні мережі використовуються для індивідуального та колективного придбання, організації, управління та використання знань. Знання отримуються з семантичних мереж, виконуючи міркування та умовиводи щодо даних мережі. Знання, що зберігаються та виводяться із смислових мереж, не повинні бути фактично правильними чи логічними. Це пов'язано з двома причинами: По-перше, семантичні мережі були розроблені для відображення того, що розуміється під інформацією. Це значення може відрізнятися від істинного умовного змісту або найбільш ймовірного тлумачення якоїсь інформації. По-друге, залежно від даних, з яких була

побудована мережа, семантична мережа може представляти універсальні, культурно-залежні, доменні та індивідуальні знання.

Логічна модель представлення знань. Логічна (предикатна) модель представлення знань заснована на алгебрі висловлювань та предикатів, на системі аксіом цієї алгебри та її правилах виводу. З предикатних моделей найбільше поширення отримала модель предикатів першого порядку, що базується на термах та предикатах. Всі знання про предметну область в такій моделі описуються у вигляді формул такої логіки або правил виводу [4]. Опис у вигляді формул дає можливість представити декларативні знання, а правила виводу – процедурні знання. Модель предметної області за допомогою логічної моделі можна визначити спрощено у вигляді: <Модель предметної області> = <понятійні знання> + <конструктивні знання>. Логічні моделі зручні для представлення логічних взаємозв'язків між фактами, вони формалізовані, для їх використання є зручний і адекватний інструментарій, наприклад, мова логічного програмування Пролог [5]. В основі моделей такого типу лежить формальна система, що задається четвіркою виду:

$$M = \langle T, P, A, B \rangle,$$

де множина T є множиною базових елементів різної природи.

множина P є множиною синтаксичних правил.

множина A є множиною аксіом.

множина B є множиною правил виведення.

Логіка була розроблена як спроба створити універсальну мову на основі математичних принципів. Отже, він заснований на формальних принципах, які представляють певні вимоги до мови представлення знань як логіку:

- Лексика - це сукупність символів, представлених символами, словами, іконками або навіть звуками. Ці символи поділяються на чотири групи:
- Логічні символи: вони не залежать від домену, наприклад кількісні показники типу "V" або сполучники типу "Λ".

- **Константи:** це залежні від домену і ідентифікують людей, властивості або відносини в області застосування або всесвіті дискурсу.
- **Змінні:** вони є необмеженими символами, область застосування яких регулюється кванторами.
- **Знаки пунктуації:** це утилітні символи, які розділяють або групують інші символи, наприклад. коми і дужки.
- **Синтаксис:** логіка повинна мати граматичні правила, які визначають, як символи поєднуються для формування добре сформованих речень.
- **Семантика:** потрібно робити змістовні твердження. Вона включає теорію відліку, яка визначає, як константи і змінні ставляться до речей у Всесвіті дискурсу. Більше того, вона також включає теорію істини для відмежування правдивих тверджень від помилкових.
- **Висновок:** цей аспект важливий для того, щоб отримати щось більше, ніж позначення. Висновок здійснюється за правилами, що визначають, як формуються зразки з інших. Відповідні правила висновку дозволяють автоматизувати механізми міркування і, таким чином, генерувати нові знання з попередніх.

Природні мови можуть представляти більш широкий спектр знань, однак логіка дозволяє виражати точно сформульовану підмножину в обчислювальній формі. З іншого боку, хоча існують деякі види знань, не виразні логікою, такі знання не можуть бути представлені ні на будь-якому цифровому комп'ютері, ні в будь-якому іншому позначенні. Виразна сила логіки включає в себе всі види інформації, що зберігається або запрограмована на будь-якому цифровому комп'ютері.

Існує багато типів логіки; кожен з них особливо підходить до своїх цільових областей застосування та має виражені можливості та вимоги до

обчислень. Логічні системи варіюються в шести вимірах від того, що можна вважати опорною логікою, логікою першого порядку.

Синтаксис: найочевидніша, але найменш важлива різниця між логіками - це позначення. З точки зору експресивної сили синтаксичні відмінності неважливі.

- Введена логіка першого порядку - це синтаксичне розширення логіки першого порядку, має однакову семантику і існують прямі синтаксичні підстановки для перекладу між ними:

$$(\forall x:t)\varphi(x) \equiv (\forall x)(t(x) \rightarrow \varphi(x)) \text{ і } (\exists x:t)\varphi(x) \equiv (\exists x)(t(x) \wedge \varphi(x)).$$

Оператори: кожна логіка визначає набір допустимих операторів або їх комбінації.

- Логіка першого порядку має загальні булеві оператори: кон'юнкція (\wedge), диз'юнкція (\vee), заперечення (\neg), імплікація (\rightarrow) та еквівалентність (\equiv), плюс універсальні (\forall) та екзистенціальні (\exists) квантори.
- Логіка ріжкових застережень є підмножиною логіки першого порядку, вона не має диз'юнкції (\vee) в імплікаційних (\rightarrow) висновках, це справа від імплікації.
- Пропозиційна логіка також є підмножиною логіки першого порядку; він включає булеві оператори, але немає кількісних показників.

Теорія доказів: інший вид змін - обмеження або розширення дозволених доказів.

- Лінійна логіка обмежує докази, дозволяючи кожним пропозиціям використовуватись лише один раз у доведенні.
- Немонотонна логіка, натомість, розширює процедури доказування, вносячи припущення за замовчуванням, якщо вони відповідають тому, що відомо в даний час.

Теорія моделі: вона визначає, як логіка пов'язана зі світом, тобто позначення або значення істинності логічних тверджень.

- Логіка першого порядку має два значення для характеристики цього відношення, істинне чи помилкове.
- Логіка, що оцінює дерево, вводить третє значення, невідоме, яке характеризує твердження, позначення яких неможливо визначити.
- Нечітка логіка багатозначна з характеристиками, коефіцієнтами визначеності, від 0,0, достовірність достовірності, до 1,0, визначеність хибною.

Онтологія: Інтерпретована логіка не має заздалегідь визначених предикатів для представлення будь-якого предмета; єдиними його символами є квантори, булеві оператори та змінні. На практиці для побудови надаються певні блоки, що полегшують їх використання, деякі типи логік включають заздалегідь визначені предикати і аксіоми, їх вбудовані онтології.

- Теорія множин використовується для забезпечення математичних основ.
- Тимчасова та динамічна логіка забезпечують часові онтології.

Метамовність: це мова про мову. Мета-мови використовуються для визначення, зміни або розширення будь-якої мови.

- Логіку першого порядку можна використовувати як метамову для будь-якої версії логіки, включаючи саму себе.
- Контекстна граматика - це версія пункту «Ріг», що використовується як метамовля для визначення синтаксису мов. Загалом, граматики еквівалентні деякій підмножині логіки першого порядку, що використовується як метамовність.
- Модальна логіка - це розширення метамовлення логіки першого порядку для розміщення модальних допоміжних дієслів, яке не використовується для розмови про те, яким є світи, але таким, яким воно може, може, повинно, повинно, могло б бути чи могло бути.

Основна модальна логіка передбачає два режими:

- р обов'язково вірно (обов'язково): $\Box p \equiv \neg \Diamond \neg p$.

- p можливо, правда (може): $\Diamond p \equiv \neg \Box \neg p$.

Подання знань продукційними правилами. Продукційна модель – це модель, заснована на правилах, яка дозволяє представити знання у вигляді пропозицій типу «Якщо (умова), то (дія)» [6]. Під "умовою" (антицедентом) розуміється деяка пропозиція – зразок, за яким здійснюється пошук у базі знань, а під "дією" (консеквентом) – дії, що виконуються при успішному результаті пошуку (вони можуть бути проміжними, виступаючими далі як умови, і термінальними або цільовими, завершальними для роботи системи). Найчастіше всього висновок на такій базі знань буває прямий (від даних до пошуку мети) або зворотний (від мети для її підтвердження - до даних). Дані – це вихідні факти, що зберігаються в базі фактів, на підставі яких запускається машина виводу або інтерпретатор правил, що перебирає правила з продукційної бази знань [7].

Простота і наочність цього способу зумовили його застосування в багатьох системах. Системи обробки знань, що використовують уявлення знань продукційними правилами, отримали назву продукційних систем. До складу продукційної системи входить база даних, база правил і інтерпретатор правил.

База правил - це область пам'яті, яка містить базу знань - сукупність знань, представлених у формі правил виду «ЯКЩО ... ТО»; база даних - це область пам'яті, що містить фактичні дані (Факти), які описують дані, що вводяться і стану систем.

Бази даних у різних системах мають різну форму, проте, всі вони можуть бути описані як група даних, що містять ім'я даних, атрибути і значення атрибутів. Інтерпретатор представляє собою механізм виведення, він є тим компонентом системи, який формує укладення, використовуючи базу правил і базу даних.

Розглянемо висновки, засновані на продукційних правилах.

Механізм, реалізований сьогодні як засіб висновків в продукційній системі, не є складним. Він має функції пошуку в БЗ, послідовного виконання операцій над знаннями і отримання висновків.

Причому існує два способи проведення таких висновків:

- прямі висновки і зворотні висновки. У прямих висновках здійснюється

просування до поставленої мети з послідовним застосуванням правил до даних (фактів), які приймаються за відправну точку.

У прямих висновках обирається один з елементів даних, що містяться в БД, і якщо при зіставленні цей елемент узгоджується з посилкою правила, то з правила виводиться відповідний висновок і поміщується в БД, або ж виконується дія, яке визначається правилом, і відповідним чином змінюється вміст бази даних.

Найчастіше такі висновки називають висновками, керованими даними, або висхідними висновками, коли послідовно виводяться нові результати, починаючи з уже відомих даних.

Висновки, при яких процес рухається в напрямку від поставленої мети до відправної точки, є зворотними. Вони називаються також такими, що сходять, або висновками, що орієнтовані на мету.

Процес низхідних висновків починається від поставленої мети. Якщо ця мета узгоджується з висновком правила, то посилка правила приймається за підціль або гіпотезу. Цей процес триває до тих пір, поки не буде отримано збіг підцільі або гіпотези з отриманими даними.

Не можна категорично відповісти на питання, який же з цих способів краще, оскільки це залежить від тієї проблеми, для якої вони використовуються. У систем, до яких ставиться вимога високої універсальності, необхідна наявність обох способів виведення.

Зазвичай в знаннях, представлених продукційними правилами, одночасно присутні кілька умов в посилці, проте є одне єдиний висновок. Вимоги до механізму висновків визначаються з умови узгодження знань,

використовуваних для висновків, причому обробка стане особливо легкою, якщо висновок обходиться одноразовим узгодженням. У цьому сенсі спадні висновки є простими висновками, однак, існують завдання, для яких не так просто дати чіткий опис мети. Наприклад, в системах медичної діагностики, коли ставиться діагноз за симптомами, важко отримувати спадний висновок, якщо невідома мета.

Онтологічна модель. У дослідженні онтологій питання виникають з перших кроків. Так, до цих пір не існує єдиного визначення для поняття онтологія. Саме поняття онтологія походить від грец. «Онтос» - суще, «логос» - вчення, поняття, тобто це розділ філософії, що вивчає буття.

Вікіпедія визначає онтологію (в інформатиці) як спробу всеосяжної і детальної формалізації деякої області знань за допомогою концептуальної схеми [8]. Під концептуальною схемою мається на увазі *набір понять + інформація про поняття* (властивості, відносини, обмеження, аксіоми і твердження про поняттях, необхідних для опису процесів вирішення завдань в обраній предметній області).

Серед фахівців, що займаються проблемами комп'ютерної лінгвістики, найбільш усталеним (класичним) вважається визначення онтології, дане Губертом: «*онтологія* - це специфікація концептуалізації» [9].

Так само відомий ряд розширених визначень Губерта, серед яких можна

виділити такі:

- *онтологія* - це експліцитно специфікація концептуалізації, де в якості концептуалізації виступає опис безлічі об'єктів предметної області та зв'язків між ними;

- *онтологія* - це знання, формально представлені на базі концептуалізації.

Формально онтологія складається з термінів, організованих в таксономії, їх визначень і атрибутів, а також пов'язаних з ними аксіом і правил виведення;

- *онтологія* - формальна специфікація розділяється концептуалізації, яка

має місце в деякому контексті предметної області;

- онтологія - база знань, що описує факти, які передбачаються завжди істинними в рамках певного суспільства на основі загальноприйнятого сенсу використовуваного словника. [10]

Розглянемо формальний опис онтології, яке вважається класичним:

Онтологія визначається як:

$$O = \langle X, R, F \rangle, \text{ де}$$

- X - кінцеве безліч понять предметної області;
- R - кінцеве безліч відносин між поняттями;
- F - кінцеве безліч функцій інтерпретації [11].

Однак далеко не всі існуючі сьогодні онтологічні ресурси підпадають під наведене визначення. Сьогодні еволюція прикладних інформаційних систем йде в бік підвищення їх інтелектуальності. Це робить істотний вплив на напрямки наукових і технологічних досліджень, пов'язаних з використанням комп'ютерів, а також надає суспільству практично значущі результати. Однак на певному щаблі розвитку подальший розвиток технологій наявними засобами стає неможливим. У такі періоди потрібен якісний стрибок використовуваних в розробці засобів. Одним з таких стрибків в області ІІІ, спрямованим на подальшу інтелектуалізацію систем взаємодії з користувачем, і стала поява онтологій.

Оскільки онтології були відповіддю науки на потреби свого часу, то і поява їх відбулося відразу в декількох областях знання. Відповідно, в кожній з них ресурси онтологічного типу були сформовані за своїми, специфічним для галузі знань, правилами.

В проектуванні онтологій умовно можна виділити два напрямки, до деякого часу розвивалися окремо. [12]

Перший напрямок пов'язаний з поданням онтології як формальної системи, заснованої на математично точних аксіомах (тобто це ресурси онтологічного типу, створені в різних областях математики).

Другий напрямок розвивався в рамках комп'ютерної лінгвістики та когнітивної науки. Тут онтологія розумілася як система абстрактних понять, які існують лише в свідомості людини, яка може бути виражена на природному мовою (або засобами якоїсь іншої системи символів). При цьому зазвичай не робиться припущень про точність або несуперечності такої системи.

Таким чином, існує два альтернативних підходи до створення і дослідження онтологій. Перший (формальний) - заснований на логіці (предикатів першого порядку, дескриптивної, модальної і т.д.). Другий (лінгвістичний) - заснований на вивченні природної мови (зокрема, семантики) і побудові онтологій на великих текстових масивах, так званих корпусах.

Онтології вказують набір елементів і відносин між елементами, які можливі або допустимі в даній області. Наприклад, створюючи структуровані подання даних про події, онтологія може приймати поняття, які представляють, хто, що, коли, де, чому і як подія. Відображення від точок даних, як правило, слів, до таких концепційних класів може бути заздалегідь задано у словниках або тезаурі, або їх можна ідентифікувати з даних. Ці відносини часто впливають із:

- Синтаксис або граматики основних даних, таких як вищезгадані трійки предметів, предикати та об'єкти.

- Відповідні відносини між поняттями, сприйняті аналітиком, які, наприклад, використовуються в методології ґрунтованої теорії.

Онтології також можуть бути ієрархічно структуровані і можуть спричинити успадкування функцій від батьківських вузлів до дочірніх вузлів. Прикладом ієрархічної онтології з успадкуванням є філогенетичні дерева, також відомі як еволюційні дерева. Співвідношення між елементами онтології можуть бути:

- Структурні, наприклад, посилання з вмісту або терміна покажчика на шматок тексту або покажчики між веб-сторінками.

- Логічні, такі як відносини еквівалентності ("є") і підтипові відносини ("є частиною від "). Область штучного інтелекту передбачає логіку, яка може бути використана для моделювання певних взаємозв'язків між поняттями

1.2. Порівняльний аналіз моделей представлення інформації.

Проаналізуємо існуючі моделі представлення знань, які використовуються саме для вирішення задачі представлення знань предметної області, наведемо переваги та недоліки кожної з моделей.

Продукційні моделі представлення знань. Продукції є найбільш популярними засобами представлення знань в інформаційних системах. Будь-яке правило продукційної моделі складається з однієї або кількох пар атрибут-значення. У робочій пам'яті продукційної системи зберігаються пари атрибут-значення, істинність яких встановлена в процесі виконання конкретного завдання. Вміст робочої пам'яті змінюється в процесі виконання завдання [13]. При описі знань конкретної предметної області для структуризації ресурсів окрему сутність предметної області розглядають як об'єкт, а дані, що зберігаються в робочій пам'яті, представляють значення, які приймають атрибути цього об'єкта. Однією з переваг такого представлення знань є уточнення контексту, в якому застосовуються правила. Правила з бази правил можуть спрацьовувати більше одного разу в процесі одного логічного висновку, оскільки одне правило може застосовуватися до різних екземплярів об'єкта (але не більше одного разу до кожного екземпляру).

Продукційна модель найчастіше застосовується для виробничих предметних областей. Вона характеризується наочністю, високою модульністю, легкістю внесення доповнень і змін і простотою механізму логічного виведення.

Фреймові моделі представлення знань. Фрейм – це структура даних для представлення деякого концептуального об'єкта [14].

У цілому фреймова модель допускає відношення всіх властивостей декларативних і процедурних знань.

Переваги методу представлення знань на основі фрейму описано нижче:

- 1) Представлення фрейм-знань робить розробку більш простим шляхом групування пов'язаних даних;
- 2) Порівняно з іншими методами подання знань, описаний у формі правила розробки, фрейм є гнучким та інтуїтивним у багатьох сферах застосування;
- 3) Представлення фрейму легко зрозуміти і використовувати людям, які не є розробниками моделі;
- 4) Не важко додати слоти для нових атрибутів та відносин;
- 5) Включити дані за замовчуванням і легко виявити відсутні значення.

Фреймове представлення знань має деякі недоліки, описані нижче:

- 1) Використовувати фреймову систему в програмі важко, тому алгоритм необхідний в процесі використання фрейму в програмі;
- 2) відсутність комп'ютерного програмного забезпечення з низькими цінами;
- 3) Механізм умовиводу не легко обробляється в системі фрейму.

Логічні моделі представлення знань. Логічне представлення - це мова з деякими конкретними правилами, яка стосується пропозицій і не має неоднозначності в представленні. Логічне уявлення означає робити висновок на основі різних умов. Це представлення визначає деякі важливі правила спілкування. Він складається з точно визначеного синтаксису та семантики, яка підтримує звуковий висновок. Кожне речення можна перевести в логіку, використовуючи синтаксис та семантику.

Синтаксис:

- Синтаксиси - це правила, які визначають, як ми можемо побудувати юридичні речення за логікою.
- Він визначає, який символ ми можемо використовувати у поданні знань.
- Як написати ці символи.

Семантика:

- Семантика - це правила, за допомогою яких ми можемо інтерпретувати речення в логіці.
- Семантичне також передбачає присвоєння значення кожному реченню.
- Логічне представлення можна класифікувати в основному на дві логіки:
 - Логіка пропозицій
 - Логіка предикатів

Переваги логічного подання:

1. Логічне представлення дозволяє нам робити логічні міркування.
2. Логічне подання є основою для мов програмування.

Недоліки логічного подання:

1. Логічні уявлення мають деякі обмеження і з ними складно працювати.
2. Техніка логічного представлення може бути не надто природною, а умовивід може бути не таким ефективним.

Мережні моделі. Можна зауважити, що для семантичної мережі, як правило, притаманне чудове представництво, а також значна потужність, що пояснює, чому багато людей складають сильний і адаптивний підхід до представлення знань. Семантичні мережі мають деякі переваги, наведені нижче:

- 1) незважаючи на різноманітність сутностей, вони можуть бути показані в одній смисловій мережі;
- 2) семантичні системи подають графічний вигляд з проблемного місця, і з цієї причини вони можуть бути простими в реалізації та легкими для розуміння;

3) Семантична мережа може використовуватися як типовий додаток для зв'язку між різними галузі знань, наприклад для інформаційних технологій;

4) семантична мережа дозволяє простий підхід до дослідження проблемного простору;

5) семантична мережа дає підхід до створення гілок споріднених компонентів;

6) семантична мережа відбивається на методах, якими люди обробляють дані;

7) смислова мережа є більш природним, ніж логічне подання;

8) семантична мережа характеризується більшою когнітивною адекватністю порівняно з формалізмом, заснованим на логіці;

9) семантична мережа дозволяє використовувати ефективний алгоритм умовиводу (графічний алгоритм);

10) Семантична мережа має більшу виразність порівняно з логікою.

Семантична мережа також має ряд недоліків, які часто викликають проблеми. Нижче наведено деякі недоліки:

1) Не існує різниці між окремими людьми та класами. Система обмежена знаннями користувача про визначення з посиланнями в семантичній мережі. Посилання між вузлами не найбільш схожі на функції. Потрібно розрізнити ланки, що містять ряд зв'язків, і ланки, які мають структурний характер. Ті ж посилання можуть використовуватися для з'єднання трьох вузлів, щоб показати структуру мережі.

2) Різниця між ознаками, що відносяться до класу, та ознаками походить від особистості та від класу, який не існує;

3) Звичайна семантика насправді не існує; тому не існує узгодженої ідеї того, що вказує запропонований дизайн зображення Семантичні системи зазвичай базуються на техніках, що їх змінюють. Альтернативою цій проблемі може бути як використання концептуальних діаграм, так і формалізм щодо представлення знань KL-ONE, що дозволяє завоювати

семантичну невизначеність у семантичній системі. KL-ONE - популярна система представлення знань у семантичній мережі та фреймі.

Онтологічні моделі. Одним із існуючих на сьогоднішній день підходів до ідентифікації предметної області, заснованих на ідеї концептуального моделювання, є онтологічне моделювання. Концептуальна або понятійна модель предметної області описує її як сукупність понять (концептів, термінів) і відношень між ними, яким відповідають сутності з реального світу [15, 16]. Цьому відповідає класичне представлення онтологічної моделі, в якому онтологія задається трьома кінцевими підмножинами: концептів, зв'язків і функцій інтерпретації. При моделюванні предметної області як сфери діяльності, відношення між поняттями також є поняттям, що описують відношення [17]. Поняття, віднесені до класу відношень, використовуються для опису процесів і явищ реального світу. Змістовна модель предметної області для понятійної моделі задається орієнтованим поміченим графом, вершини якого інтерпретуються як інформаційні елементи, що відповідають реальним об'єктам предметної області. Виділяються два типи відношень в об'єднанні моделей: змістовні, що визначають відношення одного інформаційного елемента до іншого, і понятійні, що визначають відношення елемента до концепту з понятійної ПО [18].

Онтологічний підхід дозволяє проводити постійне удосконалення моделі на основі базових онтологій шляхом їх добудови та розвитку [19]. Онтологія включає як опис предметної області, так і опис релевантних їй ресурсів. Частина онтології, що описує конкретну предметну область, включає в себе сукупність термінів і відношень семантично значущих для даної предметної області, а також правил, згідно з якими можна будувати твердження про елементи предметної області [20].

Проведений аналіз дозволяє стверджувати, що вибір виду моделі представлення знань залежить від вимог які висуваються до проектованої системи. Так, з розглянутих моделей, продукційні доцільно використовувати в випадку розробки системи направленої на реалізацію експертних висновків

з жорсткою структурою «причина-наслідок»; фреймові моделі використовуються для реалізації сценарного підходу з метою створення систем орієнтованих на прогнозування та моделювання майбутніх подій; логічні моделі застосовуються для представлення інформації у вигляді формальних математичних структур, які доцільно використовувати за необхідності оптимізації чи автоматизації певних процесів під час проектування інформаційних систем; онтологічні моделі можливо застосовувати як з метою представлення інформації в процесі проектування системи, так і для подальшого оперування такою інформацією. Саме онтологічні моделі можна використовувати як універсальний засіб для інтеграції даних та знань із різномірних джерел інформації. Використання онтологій дозволяє структурувати, систематизувати та класифікувати інформацію. Важливою характеристикою онтологічних моделей представлення знань є те, що структура онтологічної моделі апіорі вписується в парадигму семантичної павутини, що робить можливим автоматизоване налаштування та використання розробленої онтології в Інтернет середовищі. Така характеристика є затребуваною з точки зору розробки сучасних інформаційних систем, основною вимогою до яких є надання он-лайн доступу користувачам у режимі реального часу. Ще однією особливістю онтологічних моделей є можливість проведення логічних висновків, за рахунок наявних в моделі відношень асоціативного виду. Логічні висновки дозволяють організовувати змістовний пошук по інформаційному простору моделі, а також проводити певні заключення (умовиводи) по знанням, які зберігаються в моделі. Саме така особливість може бути покладена в основу оцінки якості показників певної предметної області, якщо данні та знання такої предметної області представити за допомогою онтологій.

На основі вищезазначеної інформації можна розглянути порівняння властивостей основних моделей представлення знань.

Вимоги до моделей представлення знань	Логічна модель	Продукційна модель	Фреймова модель	Мережна модель	Онтологічне моделювання
Представлення знань з точки зору природної мови	±	+	+	+	+
Декларативне представлення знань	+	+	+	+	+
Процедурне представлення знань	+	+	+	-	±
Представлення логічних зв'язків у домені	+	+	-	-	+
Представлення смислових відносин у галузі	+	±	+	+	+
Наочність опису знань	-	-	-	+	+
Цілісність представлення структури знань	-	-	-	+	+

Рис. 1.1 Порівняння властивостей основних моделей представлення знань

Дані, представлені в на рисунку (символ "+" вказує на наявність відповідних властивостей у моделі представлення, символ "±" - часткова присутність, символ "-" - відсутність) показують, що онтологічне моделювання відповідає найбільшій кількості вимог. Як результат, моделювання, яке повинне бути розроблене, має базуватися на цій моделі з урахуванням розвитку її логічних та процедурних властивостей.

Висновки

1. Проведений загальний огляд моделей представлення знань, було виявлено основні характеристики кожної із моделей представлення знань.
2. Проведений порівняльний аналіз методів представлення знань дозволив визначити доцільність застосування онтологічної моделі для моделювання різномірної слабоструктурованої інформації для визначення показників діяльності наукових установ.

РОЗДІЛ 2.

ПРЕДСТАВЛЕННЯ РІЗНОРІДНОЇ СЛАБОСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ НАУКОВИХ УСТАНОВ

2.1. Характерні особливості інформації наукових установ

В процесі організації наукової діяльності, з послідуною її оцінкою, задіяні декілька основних процесів: організація діяльності функціонування установи; визначення показників, за якими можна виявити рівень виконання певного виду діяльності в науковій установі; наявність критеріїв оцінки таких показників, за допомогою яких можна виконати якісну оцінку досягнень організації та встановити чи відповідає її діяльність вимогам, які висуваються; організація безпосередньо процесу оцінювання установи з визначенням послідовності дій, обігу відповідних документів та задіяних організацій.

Можна виділити наступні показники, за якими буде оцінюватися установа:

- Документи, що стосуються організаційного устрою
- Статистичні дані щодо працівників Установи
- Кількість наукових працівників за посадами
- Кількість виконавців НДР та їх середній вік
- Кількість працівників Установи, які задіяні у виконанні НДР, за категоріями персоналу
- Кількість дослідників за статтю та науковим ступенем
- Середній вік дослідників
- Кількість працівників, задіяних у виконанні НДР (з урахуванням сумісників), за рівнем освіти
- Список членів Вченої ради Установи
- Списки членів спеціалізованих рад Установи із захисту дисертацій
- Список членів ради молодих вчених

- Національні та міжнародні наукові спільноти
- Значення та актуальність Установи для національної стратегії в галузі науки
- Регіональне значення та актуальність Установи для регіональної стратегії в галузі науки та розвитку регіону
- Загальна концепція і розвиток Установи та інформація щодо її діяльності за попередні роки
- Результати діяльності Установи в цілому за 5 років
- Дослідження
- Виконані НДР
- Список і копії найважливіших публікацій Установи
- Кількість публікацій, підготовлених співробітниками Установи за 5 років
- Науково-навчальна література
- Електронні наукові публікації
- Проблемно зорієнтовані бази даних
- Науково-популярні видання та статті
- Публікації та виступи у засобах масової інформації
- Надання наукових послуг та вирішення інфраструктурних завдань
- Наукові консультації
- Передача знань та технологій
- Результати створення та використання об'єктів права інтелектуальної власності за 5 років
- Список проектів Установи, фінансованих на конкурсній основі з національних та зарубіжних джерел
- Наукові заходи та зв'язки з громадськістю
- Стратегічне планування діяльності Установи на наступні декілька років - потенціал розвитку науково-дослідної сфери і пріоритети

- Відповідність устаткування, обладнання та кадрового забезпечення для виконання робочих планів
- Співпраця і система наукових зв'язків Установи
- Підвищення рівня кваліфікації співробітників і кар'єрне зростання молодих науковців
- Співпраця Установи з навчальними закладами

2.2. Модель представлення інформації наукових установ

Виходячи з попередньої інформації пропонується в онтології діяльності наукових установ виділити декілька онтологій. В результаті такого виділення загальна онтологія діяльності наукових установ буде являти собою систему онтологій, кожна з яких буде описувати відповідний підпроцес в рамках як діяльності самої установи, так і процесу її оцінки зовнішніми організаціями та міністерствами. Сама ж система дозволить в повній мірі описати усі необхідні елементи, процеси та їх значення для визначення рівня ефективності діяльності установи.

Система онтологій для моніторингу та оцінки діяльності наукових установ являє собою сукупність декількох компонентів (Рис. 2.1). Розроблена система включає такі онтології як:

- Онтологія діяльності наукових установ (ОДНУ) – відображає усі можливі процеси які відбуваються в рамках наукового процесу в межах діяльності наукових установ. ОДНУ описує процес функціонування та взаємодії наукових установ, об'єкти та суб'єкти які мають відношення до такого процесу, базові елементи які задіяні на всіх рівнях нагляду, моніторингу та оцінки наукового процесу який проходить в певній науковій установі.
- Онтологія діяльності є інтегральною онтологією та поєднує наступні онтології: онтологія організації наукової діяльності, онтологія

оцінювання ефективності діяльності, онтологія моніторингу наукової діяльності.

- Онтологія організації діяльності установи (ООДУ) – описує загальні поняття, які стосуються організації наукової діяльності в цілому. Онтологія організації наукової діяльності є онтологією верхнього рівня, та включає загальні поняття стосовно процесу організаційного процесу, який відбувається в установах різного типу та виду.
- Онтологія показників діяльності установи (ОПДУ) – дозволяє детально описати показники по усім аспектам діяльності установи. На базі таких показників з врахуванням критеріїв оцінки в подальшому можна проводити якісну оцінку діяльності установи.
- Онтологія оцінювання діяльності наукових установ (ООДНУ) вводить поняття, які дозволяють провести оцінку ефективності наукової діяльності певної наукової установи в рамках її державної перевірки або встановлення її наукового рівня. ООДНУ включає в себе дві онтології, описують усі необхідні для оцінки установи елементи.
- Онтологія критеріїв оцінювання діяльності установи (ОКОДУ) – задає критерії для оцінки ефективності діяльності установи. Поняття ОКОДУ є контекстно-незалежними структурами для відображення показників діяльності установи в розрізі оцінки такої діяльності. В онтології, також, визначено шкалу за якою проводиться оцінка. Можна стверджувати, що використовуючи елементи ООЕД є можливість виконати оцінку індексу якості виконаних робіт в процесі функціонування певної установи.
- Онтологія організації оцінювання діяльності установи (ОООДУ) описує сам процес проведення оцінки відповідними органами. Така онтологія включає елементи для визначення послідовності дій в процесі оцінювання, обігу відповідних документів та задіяних в процесі оцінювання організацій та підрозділів.

- Онтологія предметної області – базується на системній класифікації наукової діяльності, організаційній структурі підпорядкованості організацій, градації і класифікації документів для оцінки діяльності установ. Онтологія предметної області задає структури для безпосереднього опису наукового та організаційного процесу, який відбувається в організації.



Рис. 2.1. Система онтологій діяльності наукових установ

Екземпляри класів і відношень, визначених у онтології діяльності наукових установ, утворюють інформаційне наповнення банку даних, який містить терміни, що представляють поняття предметної області. Вихідними даними для моделі представлення знань, що характеризують предметну область, є різноманітні нормативні документи, які регламентують процес наукової діяльності, організацію такого процесу, моніторингові засоби та методику оцінювання якості його проведення. Також, у якості вихідних даних для наповнення моделі предметної області виступають різноманітні

підручники, посібники, періодика, інформаційні ресурси та інше, що містять наукову інформацію яка формує основу наукової діяльності. Така інформація є базою для формування простору знань, яким оперує наука у певній предметній області.

2.3. Елементи системи онтологій для представлення інформації наукових установ

Для прикладу наведемо опис однієї компоненти системи онтологій для моніторингу та оцінки діяльності наукових установ – Онтології організації діяльності установи. Наведемо опис елементів онтології, її класи, атрибути та відношення онтології.

Онтологія організації діяльності установ включає дев'ять класів, які описують загальні поняття, які стосуються організації діяльності установи. Ці класи пов'язані відношеннями різних типів. Різні властивості кожного поняття описуються на основі атрибутів понять і обмежень, накладених на область їх значень. Наведемо класи Онтології організації наукової діяльності наукової діяльності.

Персона. До цього класу відносяться поняття, пов'язані з суб'єктами наукової діяльності: дослідниками, співробітниками і членами організацій. Атрибутами персони є: персональні дані (прізвище, ім'я, по батькові, дата народження, телефон, адреса проживання, паспортні дані), вчений ступінь, звання, напрямки наукової діяльності, нагороди, публікації, наукові теми.

Організація. Поняття цього класу описують різні наукові організації, такі як наукові інститути, академії, університети, міністерства, співтовариства і асоціації, дослідницькі групи та інші неформальні наукові об'єднання. Атрибутами організації є: назва (повна та скорочена), місце розташування, реквізити установи.

Захід. У цей клас входять поняття, що описують науково-організаційну або науково-дослідну діяльність. До класу входять такі наукові заходи як: конференції, семінари, воркшопи, дослідницькі поїздки, проекти, програми і

т.д. До атрибутів класу «Захід» відносяться: назва, місце проведення, дата початку, дата закінчення, ступінь завершеності, тематика, назва та реквізити організатора, відповідальний за захід (як від установи, яка проводить захід, так і від певної установи яка приймає участь у заході).

Діяльність. Клас описує усі види діяльності які відбуваються в науковій організації (наприклад, такі як дослідження, проект, програма досліджень, публікація). Поняття класу «Діяльність» є сполучною ланкою між класами які описують початок та кінець процесу наукової діяльності, а також підключають учасників такої діяльності та класи предметної області, які визначають наукові аспекти такої діяльності. Таким чином, можна сказати, що клас «Діяльність» задає ланцюг понять типу: вид діяльності – учасники діяльності - методом і об'єктом дослідження з предметної області які фігурують в процесі діяльності – отриманий науковий результат (як з організаційної, так і з наукової точки зору).

Публікація. Цей клас слугує для опису різного роду публікацій: у періодичних виданнях і тих, які видаються в результаті проведення конференцій, наукових заходів і т. д.). До атрибутів публікації належать: назва, автори, короткий опис (анотація, ключові слова), дата публікації та мова публікації, характеристики журналу у якому розміщена публікація (назва, вид, том, серія, номер, рік видання, відповідальні редактори, періодичність, показники цитованості, ISSN, URL)

Місцезнаходження. Цей клас понять дозволяє описувати географічну та адміністративно-територіальну локалізацію усіх концептів, які задані класами онтологічної моделі (об'єктів дослідження, організацій і т.д.). Атрибутами цього класу є назва місця розташування, географічний тип, адміністративно-територіальна приналежність, країна розташування, повна адреса.

Література. Клас описує літературу (представлену в друкованому або електронному форматах), яка використовується і продукується в процесі наукової діяльності (монографії, книги, підручники, статті, праці

конференцій, періодичні видання, звіти, результати експериментів та досліджень, фото та відеоматеріали, тощо). До атрибутів публікації належать: назва, короткий змістовний опис, дата та мова публікації, реквізити публікації (друкарські характеристики), вид публікації.

Документація. Цей клас служить для опису різного роду документації. Можна виділити декілька груп документації: законодавча (яка задає норми організації наукового процесу на рівні країни), стандартизована (ГОСТ, ДСТУ, ISO), внутрішня документація (організаційна документація, яка регламентує роботу на рівні однієї організації). До атрибутів документації відносяться: назва, короткий змістовний опис, дата та мова публікації, реквізити публікації (друкарські характеристики), вид публікації.

Науково-навчальні матеріали. Цей клас містить поняття які визначають наукову літературу, яка використовується в процесі наукової діяльності. До цього класу можна віднести такі поняття як: існуючі монографії, книги, підручники, довідники, навчальні посібники, мануали, репозитарії та класифікатори. До атрибутів класу відносяться: назва, опис, дата публікації та мова.

Класи онтології організації наукової діяльності пов'язані наступними асоціативними відношеннями:

«бути автором» – використовується для встановлення зв'язку між *Персоною*, яка є автором публікації, і самою *Публікацією*;

«входити у» – пов'язує клас *Організація* і *Персона* у разі, коли персона будь-яким чином відноситься до організації;

«бути учасником» – пов'язує *Захід* з *Персоною* або *Організацією*, що бере участь у даній події;

«бути організатором» – встановлює зв'язок між *Подією* і *Персоною* або *Організацією*, що є організатором події;

«наукові праці» – задає зв'язок між *Заходом* і *Публікаціями*, що висвітлюють цей захід;

«бути виданим у» – пов'язує *Публікацію* і *Організацію*, що видає наукові публікації;

«розташовуватися» – описує місце розташування *Організації*;

«описувати» – пов'язує *Публікацію* з будь-яким поняттям *Онтології знань*.

«частина діяльності» – пов'язує *Діяльність* з *Об'єктами*, *Методами* та *Результатами* дослідження.

Структурні відношення «частина-ціле» пов'язують наступні класи онтології організації наукової діяльності:

Відношенням «частина-ціле» описується зв'язок між класом *Література* і класами *Документація*, *Навчальні матеріали* та *Публікація*. Останні є підкласом класу *Література*.

«Частина-ціле» актуальне для класу *Персона* і класів *Дослідники*, *Співробітники* і *Члени організацій*.

Клас *Організація* є цілим для *Організацій*, *Наукових спільнот* і *Асоціацій*, *Інститутів*, *Дослідницьких груп* та інших об'єднань.

Відношенням «частина-ціле» зв'язується клас *Захід* з *Конференціями*, *Дослідницькими поїздками*, *Проектами*, *Програмами*.

Зв'язок класу *Документація* з документами типу *ГОСТ*, *ДСТУ*, *ISO* та ін. здійснюється за допомогою відношення «частина-ціле».

Відношення «частина-ціле» описує зв'язок класу *Навчальні матеріали* з *Підручниками*, *Довідниками*, *Навчальними посібниками*, *Мануалами*.

Відношення наслідування. Відношення наслідування використовується для відображення наслідування атрибутів дочірніми класами від батьківських. Для класів ООНД які мають підкласи відношення наслідування передає всі їх атрибути до дочірніх підкласів.

Відношення виду «клас-дані» використовується для зв'язку конкретних екземплярів понять з класом. Описане відношення актуально для всіх класів даної онтології.

Інші онтології система онтологій для моніторингу та оцінки діяльності наукових установ описуються аналогічно.

Онтологія дозволяє структурувати та систематизувати інформацію накопичену науковими установами для її подальшого використання в процесі оцінки ефективності функціонування установи. Загальну онтологію моніторингу та оцінки діяльності наукових установ представлено системою взаємозв'язаних компонент, кожна з яких дозволяє описати відповідний підпроцес в рамках як діяльності самої установи, так і процесу її оцінки зовнішніми організаціями та міністерствами. Сама ж система дозволить в повній мірі описати усі необхідні елементи, процеси та їх значення для визначення рівня ефективності діяльності установи. Проведено опис елементів усіх компонент системи онтологій моніторингу та оцінки діяльності наукових установ.

Висновки

1. Розглянуто та наведено основні характерні особливості інформації наукових установ.
2. Запропонована модель представлення різномірної слабоструктурованої інформації.
3. Описано системи онтології для представлення різномірної слабоструктурованої інформації, яка дозволить в повній мірі описати усі необхідні елементи, процеси та їх значення для визначення рівня ефективності діяльності установи .

РОЗДІЛ 3.

ОНТОЛОГІЯ РІЗНОРІДНОЇ СЛАБОСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ НАУКОВИХ УСТАНОВ

3.1. Програмні середовища проектування онтологічних моделей

Для того щоб користувачам в інженерії і підтримці онтологій було легше будувати та підтримувати онтологічні моделі було розроблено ряд інструментальних засобів.

WebOnto [21] був розроблений Knowledge Media Institute of the Open University. WebOnto призначений для підтримки спільного використання, створення та редагування онтологій. Він був розроблений для доповнення Tadzebao, інструменту дослідження онтології. WebOnto – це в основному графічний інструмент для побудови онтологій. Він використовує OCML (Мова оперативного концептуального моделювання), розроблений в контексті проекту VITAL для моделювання онтологій. Інструмент має ряд корисних функцій: збереження структурних діаграм, окремий вигляд відносин, класів, правил тощо. До інших можливостей можна віднести, наприклад, спільну роботу над онтологією, використання діаграм та функцій передачі та прийому.

Protege [22] – це вільно доступна локальна програма на базі Windows, розроблена командою з медичної інформатики Стенфордського університету. Програма призначена для побудови онтологій моделей областей додатків. Її мета – допомогти розробникам програмного забезпечення у створенні та підтримці явних моделей тематичної області та включенні цих моделей безпосередньо в програмний код. Протеже складається з трьох основних частин, які необхідно використовувати по черзі. Спочатку йде редактор онтології, який дозволяє проектувати онтології, розгортаючи ієрархічну структуру і включаючи абстрактні або конкретні класи та слоти. На основі встановленої онтології Protege здатний генерувати інструмент отримання знань для введення екземплярів онтології. Остання частина програми – це

інтерпретатор схеми, який дозволяє інстанціювати для класів та підкласів. Інструмент має повний графічний інтерфейс, що дуже зручно для недосвідчених користувачів.

OntoSaurus [23] має веб-браузер для базових знань LOOM, який використовує для них графічний інтерфейс. OntoSaurus також обмежує засоби редагування, його основна функція – перегляд онтологій.

ТОДАОС – забезпечує побудову усіх ланцюгів процесу трансдисциплінарної інтеграції: семантичний контент-аналіз текстових документів; таксономізація; виділення властивостей концептів таксономії; формування онтології задачі вибору; трансдисциплінарна інтеграція контекстів на основі властивостей-критеріїв концептів, які визначають онтологію вибору; включення документів, знайдених у глобальному середовищі. За рахунок активних станів гіпервідношення множинною часткової впорядкованості [24, 25], ТОДАОС являє собою інноваційну ІТ-технологію онтологічного управління знаннями та інформаційними ресурсами, незалежна від стандартів їх створення. В основі функціональності ІТ-ТОДАОС лежать методи обробки інформації – Big Data, Data Mining, Semantic Web [26, 27]. Трансдисциплінарне управління знаннями базується на механізмах виділення з інформаційних масивів термінополів та їх таксономізації

Таким чином операціонально ІТ– ТОДАОС забезпечує:

- мережеву взаємодію з неструктурованою і слабо структурованою інформацією великих обсягів; лінгвістично-семантичний аналіз контенту, агрегування та рейтингування інформаційних ресурсів;
- багато-індексний пошук тематичної інформації з великою кількістю міждисциплінарних зв'язків та відношень;
- адаптивність під тематичний профіль діяльності кожного суб'єкта освіти на основі засобів Semantic Web інтерфейсу з мережевими інформаційними ресурсами та інтерактивними системами знань;

Галерею об'єктів онтологічного інтерфейсу та обробку онтографу (доступ до якого відбувається за допомогою MySQL, XML) реалізовано скриптовою мовою PHP, візуалізацію – HTML5, CSS, JavaScript, jQuery.

Отже, для подальшої роботи було обрано IT-ТОДАОС, детальніше про нього можна дізнатися в наступному розділі.

3.2. Характерні особливості програмної платформи ТОДАОС

Когнітивна інформаційна технологія – ТОДАОС (KIT-ТОДАОС), орієнтована на підтримку процесів семантико-лінгвістичного аналізу великої кількості просторово-розподіленої слабоструктурованої інформації, її структуризацію, встановлення контекстних зв'язку між документами, що обробляються, прогнозування та підтримка процесів раціонального вибору з наступною побудовою інформаційних та аналітичних WEB-орієнтованих рішень.

Когнітивна інформаційна технологія - ТОДАОС, орієнтовується на обробку великої кількості різноманітної текстової слабоструктурованої інформації (Big Data), в автоматичному та автоматизованому режимах, на основі семантичного та лінгвістичного аналізу, її структуризації та класифікації, підтримання процесу нормованого вибору та прогнозування з наступною побудовою інформаційних, аналітичних та мережевих рішень. Засоби KIT-ТОДАОС забезпечують реалізацію наступних інформаційних процесів:

- Лінгвістичний та семантичний аналіз мережевих інформаційних ресурсів, які мають значну кількість відношень між дисциплінами, та побудовані на основі використання різноманітних інформаційних технологій і стандартів;
- Трансдисциплінарний аналіз та інтеграція з іншими мережевими інформаційними системами та WEB-орієнтованими додатками;
- Таксономізація оповідей довільних документів та відображення їх структури, включаючи зв'язки між контекстом;

- Створення онтологічних інтерактивних документів;
- Виявлення прихованої інформації у інформаційних ресурсах, що аналізуються;
- Глибинне та машинне навчання (Deep Learning, Machine Learning);
- Підтримка форматів і протоколів Semantic Web;
- Опрацювання великих даних (Big Data).

Операційні можливості КІТ ТОДАОС:

- Засоби контекстного, семантичного та лінгвістичного аналізу природно-мовного тексту та побудова таксономії документів;
- Засоби класифікації та генерації онтологій предметної області;
- Засоби розробки трансдисциплінарної онтології;
- Онтологія задачі вибору для інформаційно-аналітичної підтримки прийняття рішень та забезпечення процесів багатокритеріального аналізу й рейтингування;
- Засоби семантичного пошуку лексичних структур на основі лінгвістичної обробки великої кількості розподілених мережних текстових масивів;
- Лінгвістичний кластер - електронна бібліотека із засобами асоціативного пошуку семантично зв'язаних інформаційних масивів.

Переваги

Забезпечуються процеси структуризації та трансдисциплінарних інтеграцій розподілених інформаційних ресурсів, процедури взаємодії мережних інформаційних систем, які мають значний обсяг міждисциплінарних відношень, та побудовані на основі використання різних інформаційних технологій і стандартах, та створення експертного середовища підтримки прийняття рішень.

З компонентною конфігурацією трансдисциплінарних процесів можна ознайомитися в таблиці 3.1

Компонентна конфігурація трансдисциплінарних процесів

Трансдисциплінарний процес	Когнітивні процедури
Автоматизоване внесення інформаційних масивів, документів з джерел тематичної інформації, визначених конфігурацією НТК ННТП ;	Внесення інформаційних масивів та документів у середовище НТК ННТП Індексне маркування інформаційних масивів Структуризація вхідної текстової інформації, що представлено українською, російською та англійською мовами Редагування таксономій Збір тематичної інформації з визначених конфігурацією НТК ННТП інформаційних джерел
Таксономічне (ієрархічне) відображення структури інформаційних масивів, що представлені українською, російською та англійською мовами;	Структуризація вхідної текстової інформації, що представлено українською, російською та англійською мовами Таксономічне відображення структури інформаційних масивів, що представлені українською, російською та англійською мовами Автоматизована індуктивна побудова таксономій Редагування таксономій Формування, редагування та видалення онтологій Автоматичне виявлення оптимальних закономірностей з різнотипних даних у явній формі логікових виразів Спільна обробка як числових так і номінальних даних у задачах виявлення та використання логікових виразів
Формування та відображення динамічних тематичних класифікаторів, каталогів та реєстрів інформаційних ресурсів;	Структуризація вхідної текстової інформації, що представлено українською, російською та англійською мовами Індексне маркування інформаційних масивів Тематична класифікація інформаційних масивів та документів Автоматичне формування динамічних тематичних класифікаторів, каталогів та реєстрів інформаційних ресурсів Агреговане представлення інформаційних ресурсів Редагування таксономій

Продовження таблиці 3.1

Індексне маркування інформаційних масивів;	Індексне маркування інформаційних масивів Структуризація вхідної текстової інформації, що представлено українською, російською та англійською мовами
Виявлення синонімічної еквівалентності між інформаційними масивами, що відображаються українською, російською, англійською, німецькою, французькою та італійською мовами;	Виявлення синонімічності між інформаційними масивами, що відображаються українською, російською, англійською, німецькою, французькою та італійською мовами Внесення інформаційних документів та масивів в індексну та синонімічну зони
Генерація онтологічних мережецентричних площадок користувачів;	Генерація онтологічних мережецентричних площадок користувачів Структуризація вхідної текстової інформації, що представлено українською, російською та англійською мовами Таксономічне відображення структури інформаційних масивів, що представлені українською, російською та англійською мовами Редагування таксономій
Формування інтерактивних документів; прогнозне оцінювання;	Формування інтерактивних документів Структуризація вхідної текстової інформації, що представлено українською, російською та англійською мовами Таксономічне відображення структури інформаційних масивів, що представлені українською, російською та англійською мовами Редагування таксономій
Підтримка процесу прийняття рішень на основі таксономічного представлення структури інформаційних масивів, що аналізуються;	Структуризація вхідної текстової інформації, що представлено українською, російською та англійською мовами Таксономічне відображення структури інформаційних масивів, що представлені українською, російською та англійською мовами Ранжування альтернатив Раціональний вибір Прогнозне оцінювання Проведення багатокритеріального порівняльного аналізу Редагування таксономій

Продовження таблиці 3.1

Формування інформаційних документів (довідок, звітів, запитів тощо)	<p>Структуризація вхідної текстової інформації, що представлено українською, російською та англійською мовами</p> <p>Таксономічне відображення структури інформаційних масивів, що представлені українською, російською та англійською мовами</p> <p>Підтримка прийняття рішень у процесах класифікації, діагностики та виявлення властивостей на основі отриманих закономірностей з використанням нечіткої логіки.</p> <p>Автоматичне виявлення оптимальних закономірностей з різнотипних даних у явній формі логікових виразів</p> <p>Реалізація спільної обробки як числових так і номінальних даних у задачах виявлення та використання логікових виразів</p> <p>Редагування таксономій</p>
---	---

Модуль перегляду онтологій забезпечує функції перегляду онтологій і пошуку в індексній зоні. Доступ до даного модуля надається при наявності ролі «Користувач», тобто всім користувачам, включаючи анонімних. При цьому анонімні користувачі можуть переглядати виключно публічні онтології (з публічних онтологічних сайтів або онтології, для яких ввімкнений доступ за посиланням), тоді як авторизовані користувачі можуть додатково переглядати онтології із свого особистого кабінету і онтології, доступ до яких надано адміністратором.

На Рис. 3.1 відмічені основні елементи інтерфейсу модуля перегляду онтологій:

- 1) Назва онтології, що переглядається;
- 2) Перемикач модулів фільтрації (з кнопкою «очистити фільтр»);
- 3) Активний модуль фільтрації;
- 4) Блок пошуку (зовнішній пошук в Google і пошук в доступних індексних зонах);
- 5) Перемикач основного режиму відображення;
- 6) Активний модуль відображення.

Справа від перемикача основного режиму відображення знаходиться кнопка відкриття головного меню (Рис. 3.1) з кнопками бібліотек і кнопкою входу.



Рис. 3.1 Стандартний модуль перегляду онтологій

Стандартний модуль відображення онтологій містить наступні режими:

- 1) Режим об'єктного відображення.
- 2) Режим табличного відображення.
- 3) Режим відображення онтографа.
- 4) Режим відображення ГІС-додатка.

Режим об'єктного відображення представляє об'єкти онтології у вигляді списків, розбитих по «рівням» (Рис. 3.2). Даний режим перегляду працює наступним чином:

- 1) Верхній рівень за-замовчуванням містить кореневі категорії і об'єкти без категорій.
- 2) Якщо ввімкнена ієрархічна фільтрація, то він складається тільки з вибраного об'єкта чи категорії, якщо ввімкнена фільтрація за атрибутами;
- 3) Користувач може вибрати категорію, натиснувши на неї, така категорія стане **актуалізованою** (1), і на її основі сформується

наступний рівень – першим елементом рівня буде сама актуалізована категорія, а іншими елементами – всі її підкатегорій і дочірні об’єкти.

- 4) Всі елементи відображаються у вигляді круглих блоків з назвою і (за наявності) зображенням, при наведенні вказівника миші блок міняє форму (2).
- 5) При наведенні на блок відображаються додаткові кнопки (3) – «Перегляд картки об’єкта» і «Ієрархічна фільтрація», що запускають відповідні функції для з асоційованою з блоком категорією чи об’єктом в якості параметра.



Рис. 3.2 Режим об’єктного відображення

Режим табличного відображення – призначений для відображення списку об’єктів з однаковим набором атрибутів (Рис.). Даний режим працює наступним чином:

- 1) Табличне відображення може відображати виключно ті об’єкти, що мають непусту множину атрибутів. Категорії з непустою множиною атрибутів в даному режимі обробляються аналогічно до об’єктів і також можуть відображатись в таблиці. Якщо онтологія не містить об’єктів з атрибутами, то перемикання на цей режим перегляду буде недоступне.
- 2) Вся множина об’єктів з атрибутами розділяється на сторінки фіксованого розміру (за-замовчуванням – 50 об’єктів). Одночасно

може відображатись тільки одна сторінка, а їх перемикання відбувається з допомогою спеціальних керуючих елементів (3).

- 3) Для поточної сторінки автоматично формується заголовок таблиці (1). Заголовки різних сторінок можуть відрізнятись. Система використовує алгоритм, що робить спробу встановити відношення порядку між атрибутами. В деяких випадках (якщо порядок слідування атрибутів в різних об'єктів різний), встановити порядок стає неможливо, тоді порядок може бути довільним.
- 4) Кожен рядок таблиці представляє об'єкт. Перша комірка рядка містить назву об'єкта (2) і його зображення (за наявності). При натисканні на назву відкривається картка об'єкта.
- 5) Деякі спеціалізовані типи атрибутів (координати, зображення) не відображаються в таблиці. Посилання відображаються спеціальним чином (4), що дозволяє користувачу робити швидкий перехід по ним.
- 6) Якщо об'єкт має кілька атрибутів з однаковою назвою, то їх значення будуть відображатись в комірці таблиці одне під одним, розділені горизонтальною лінією.

	Абсолютна твердість	Твердість (шкала Мооса)	Питома вага	Показник заломлення	Подвійне променезаломлення	Плавкість	Прозорість	Ціна	aregis_link
Самородний кобальт	1025	8.84	4.11	0.062	4.893	680.05	0.86	649.1	Карта: САМОРОДНІ ЕЛЕМЕНТИ
Самородний нікель	821	4.01	2.33	0.016	2.483	1757.3	0.55	493.3	Карта: САМОРОДНІ ЕЛЕМЕНТИ
Самородна мідь	780	7.62	2.04	0.04	2.525	388.76	0.64	959.45	Карта: САМОРОДНІ ЕЛЕМЕНТИ
Самородне срібло	915	6.58	4.69	0.025	3.758	830.22	0.58	290.02	Карта: САМОРОДНІ ЕЛЕМЕНТИ

Рис. 3.3 Режим табличного відображення

Режим відображення онтографа призначений для одночасного відображення великої кількості об'єктів і зв'язків між ними (Рис. 1)

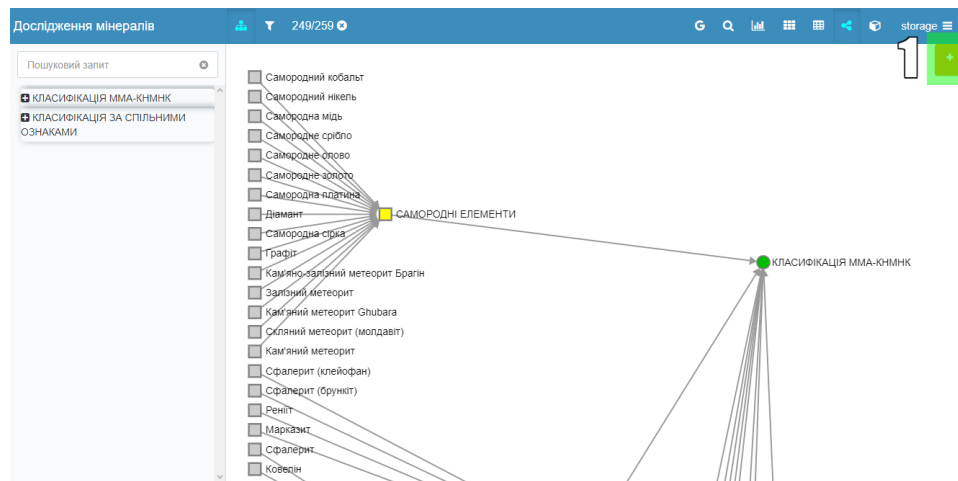


Рис. 1 Режим відображення онтографа

В режимі відображення онтографа, як і в табличному, категорії *обробляються аналогічно до об'єктів*. Об'єкти при відображенні використовують свої налаштування – форму (квадрат або круг) і колір. Також об'єкти, що мають непусту множину атрибутів, відображаються з рамкою.

Режим відображення онтографа **керується наступним чином:**

- 1) Якщо після наведення вказівника мишки на об'єкт утримувати ліву кнопку миші і переміщувати вказівник – буде здійснене перетягування (Drag&drop) відповідного об'єкта.
- 2) Якщо навести вказівник миші на пусте місце, утримувати ліву кнопку миші і переміщувати вказівник – буде здійснене переміщення робочої області (pan).
- 3) За допомогою пристрою миші дозволяє збільшувати або зменшувати робочу область (zoom).
- 4) Одинарний клік по об'єкту дозволяє виділити його, одинарний клік з натиснутим Shift – додати об'єкт до списку виділених. Виділені об'єкти можуть переміщуватись одночасно, якщо при переміщенні утримувати Shift.
- 5) Натискання комбінації клавіш Ctrl+A дозволяє вибрати всі доступні об'єкти.
- 6) Подвійний клік по об'єкту відкриє перегляд картки об'єкта (описано у відповідному розділі).

Крім того, режим відображення онтографа містить додаткові налаштування фільтрації. Відповідна панель налаштувань відкривається натисканням на кнопку (1) і детально описана в розділі «Фільтрація».

Також важливою функцією режиму перегляду онтографа є копіювання об'єктів, що здійснюється виділенням множини об'єктів натисканням комбінації Ctrl+C. Дана функція може використовуватись експертами ПГ при формуванні онтологій і детально описана в розділі «Копіювання і вставка».

Даний режим відображення має спеціалізовані інструменти ієрархічної фільтрації, показані на Рис. 2

- 1) Глибина фільтрації визначає довжину шляху між вибраним об'єктом і об'єктом, що буде показано. При цьому враховується тільки найкоротший шлях (показаний на Рис. 2 червоними стрілками), хоча при цьому можуть існувати і більш довгі (показано зеленими стрілками).
- 2) Перемикач напрямку фільтрації може мати значення «Дочірні об'єкти», «Батьківські об'єкти» і «Сусідні об'єкти». Перемикач визначає, які саме зв'язки між об'єктами будуть враховуватись при побудові множини видимих об'єктів (вхідні зв'язки, вихідні зв'язки і всі зв'язки відповідно).
- 3) Кнопка «для вибраного об'єкта» надає додатковий, більш зручний спосіб запуску процедури ієрархічної фільтрації. При натискуванні на неї для ієрархічної фільтрації буде вибраний в даний момент об'єкт. Якщо вибрано кілька об'єктів, то вибраний буде один з них довільним чином.

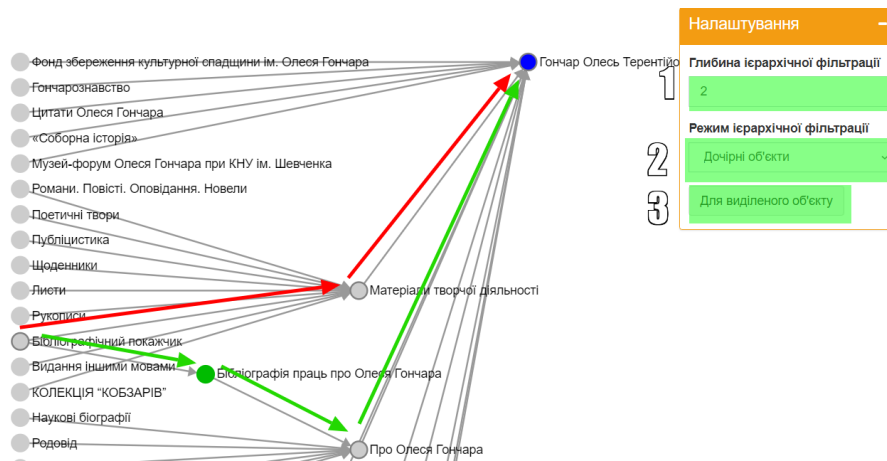


Рис. 2 Інструменти ієрархічної фільтрації в режимі онтографа

Режим онтологічної призми — призначений для відображення деревовидних структур даних (Рис. 3). Даний режим використовує три рівня (аналогічно до тих, що використовуються в об'єктному відображенні)

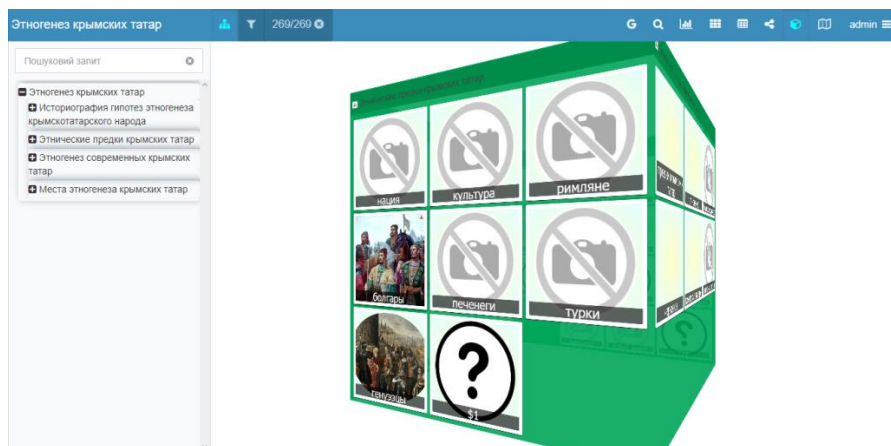


Рис. 3 Режим онтологічної призми

- 1) Перший рівень складає коренева категорія онтології. Якщо таких категорій кілька, то використовується перша з них. Така категорія представляє всю призму і явно не відображається в робочій області.
- 2) Другий рівень складють підкатегорії кореневої категорії. Вони формують грані призми, а імена цих категорії відображаються в верхній частині відповідної грані (Рис. 4, 1).
- 3) Третій рівень формується підкатегоріями і дочірніми об'єктами категорій другого рівня. Такі об'єкти представляються блоками на відповідних гранях — аналогічно до об'єктного представлення. Максимальна кількість блоків на грані — 25, якщо дочірніх об'єктів в

категорії більше, то вони розбиваються на сторінки, а в заголовку грані відображаються перемикач сторінок (Рис. 4 , 2).

- 4) Якщо через структуру онтології елементом другого рівня стає об'єкт, а не категорія, то цей об'єкт відображається і в якості грані, і в якості блоку на грані.

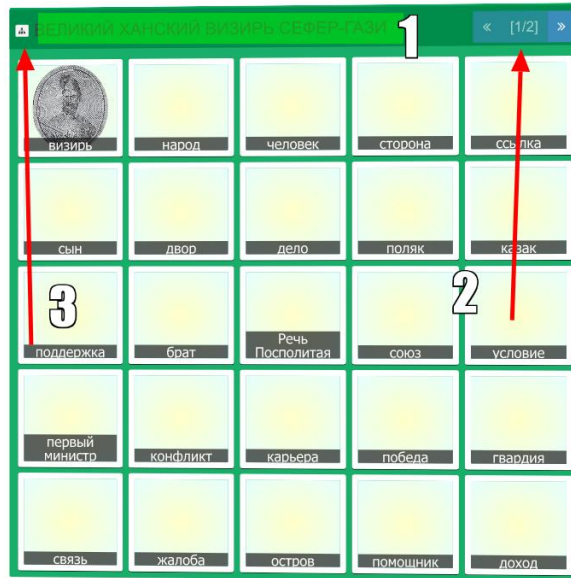


Рис. 4 Грань онтологічної призми

Для режиму відображення онтологічної призми діє специфічна, значно складніша, логіка ієрархічної фільтрації:

- 1) Якщо ієрархічна фільтрація не працює, то призма формується на основі піддерева першої (такої, що відображається найвище в режимі перегляду онтографа) кореневої категорії.
- 2) Якщо в стандартному ієрархічному фільтрі вибрана одна з кореневих категорій, то призма формується на основі її піддерева.
- 3) Якщо в стандартному ієрархічному фільтрі вибрана підкатегорія однієї з кореневих категорій, то призма формується на основі піддерева цієї кореневої категорії, при цьому при відображенні призма автоматично буде повернута так, що сформована для вибраної категорією грань була на передньому плані.
- 4) Якщо вибрано об'єкт чи категорію третього рівня (дочірній до підкатегорії кореневої категорії), призма формується на основі

кореневої категорії, сформованої підкатегорією. Відповідна грань повертається на передній план, на грані встановлюється сторінка, що містить вибраний об'єкт, і цей об'єкт виділяється кольором (Рис. 5).

- 5) Обійти описану вище складну логіку дозволяє спеціалізована кнопка (Рис. 5), що показується для всіх граней, сформованих категоріями (а не об'єктами). Кнопка дозволяє відобразити призму, сформовану даною категорією, незалежно від її місця в ієрархії.

Натискання блоку на грані призми викликає перегляд картки відповідного об'єкта.

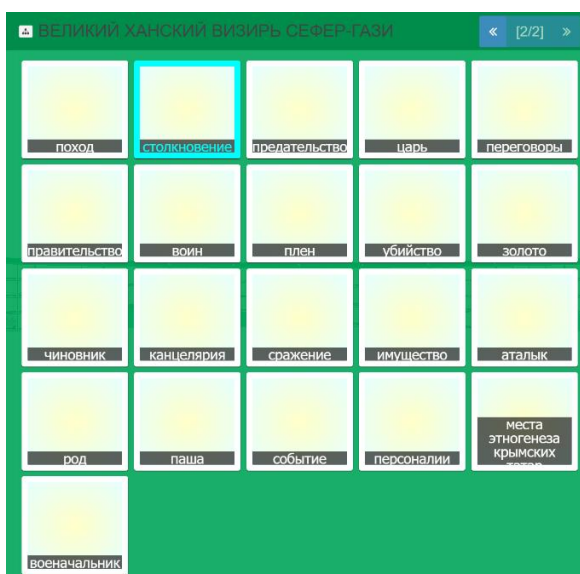


Рис. 5 Виділення об'єкта на грані

3.3. Онтологія інформації наукових установ на базі платформи ТОДАОС

Згідно запропонованого інтелектуального методу представлення інформації та обраного програмного середовища, було створено онтологічну модель представлення інформації наукової установи.

Нижче наведено скріншоти, отримані з середовища ІТ-ТОДАОС, що демонструють результати виконаної роботи.

На Рис. 3.9 відображена онтологія верхнього рівня, яка складається з:

- Рейтингу установи
- Рейтингу викладачів установи

- Рейтингу кожного викладача
- Рейтингу за певний рік

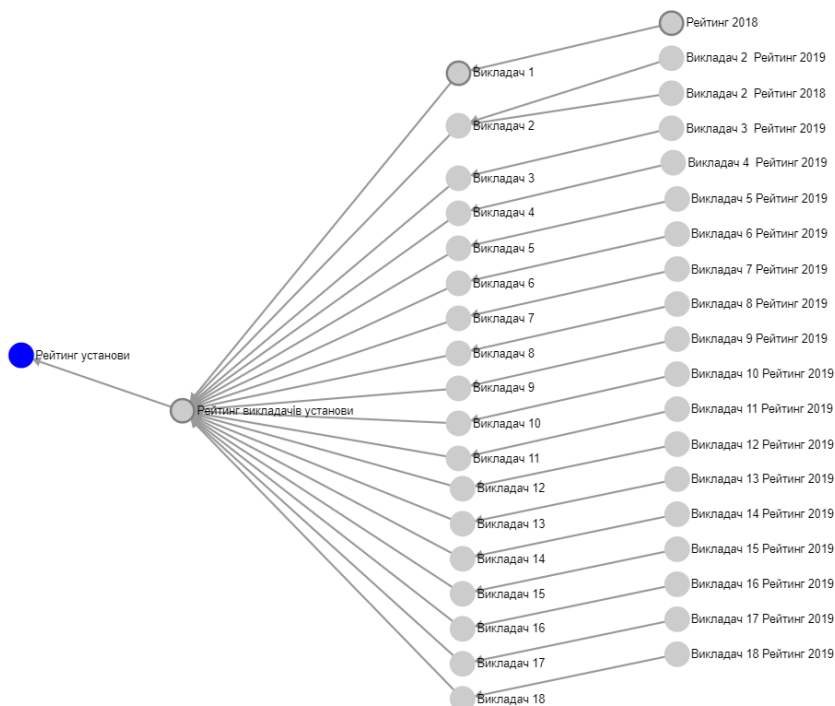


Рис. 3.9 Онтологія верхнього рівня

На Рис. 3.10 зображена онтологія нижнього рівня, на якій можна побачити, що рейтинг кожного викладача ділиться на три напрямки, а кожен напрям роботи ділиться на групи робіт.

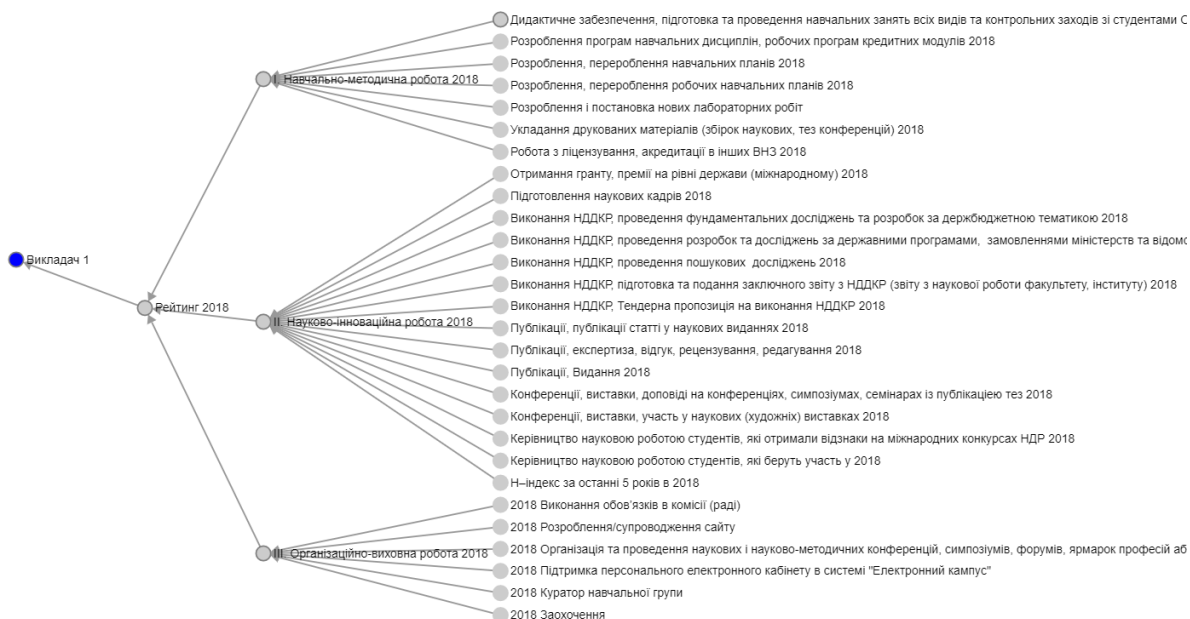


Рис. 3.10 Онтологія нижнього рівня

На Рис. 3.11 відображено, складові кожної групи робіт, а саме окремі роботи викладача.

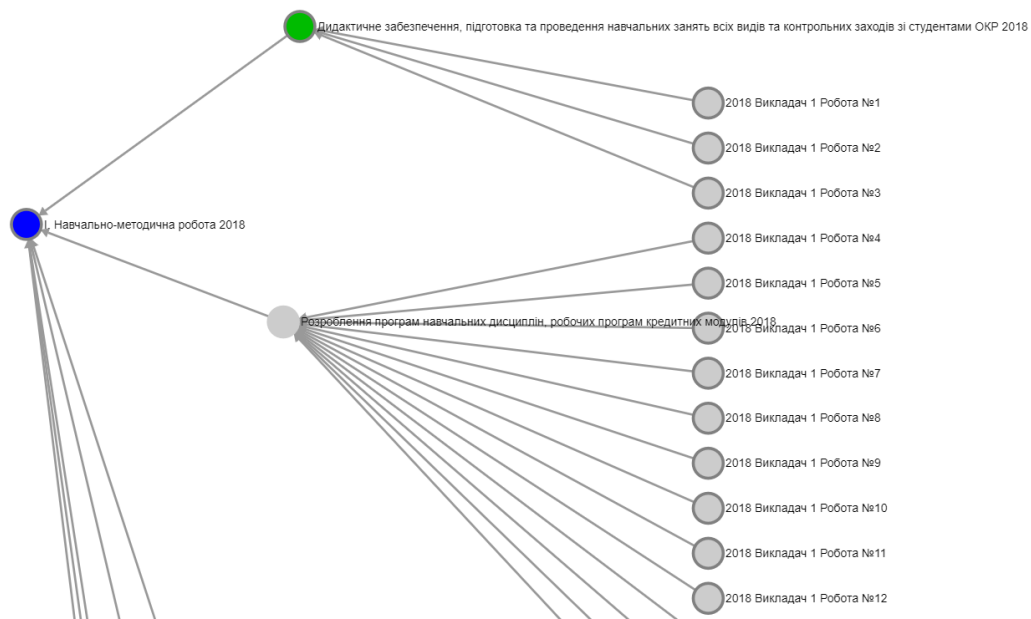


Рис 3.11 Онтолія робіт викладача

Ознаймившись із структурою онтологічної моделі установи, розглянемо складові кожного об'єкту

На Рис. 3.12 відображено атрибути об'єкту Робота №5. До атрибутів входять такі дані:

- Вид роботи
- Бали - кількість балів, яку можна отримати за роботу
- Кількість (частина виконана викладачем) у відсотках
- Результат – бали які отримав викладач за цю роботу

Рис 3.12 Атрибут об'єкту онтології

На Рис. 3.13 відображено характеристику робіт в певній групі, та загальну інформацію про роботи викладача, також кількість балів всіх робіт що входять в цю групу.

Метрика	Значення
Аудитория	аспиранты; бакалавр; магистр
Язык	Украинский язык
Часы учебного нагрузки за бюджетом	588.6
Часы учебного нагрузки за бюджетом по контракту	14.1
Общее время	602.7
Результат	783.51

Рис. 3.13 Характеристика робіт

На Рис. 3.14 відображено загальну кількість балів за всі роботи які були виконані викладачем в певному році.

Метрика	Значення
Загальний рейтинг науково-педагогічного працівника	5951.28
Загальний рейтинг кафедри	2478.75
Загальний рейтинг завідувача кафедри	6694.91

Рис. 3.14 Загальна кількість балів за роботи викладача

На рисунку 3.15 відображено загальна сума балів за всі роки.

Метрика	Значення
Загальний рейтинг викладача	15124.939999999999

Рис. 3.15 Загальна сума балів за всі роки

На Рис. 3.16 відображено загальний рейтинг всіх викладачів установи.

Метрика	Значення
Загальний рейтинг викладачів	302498.8

На Рис. 3.17 відображено фрагмент онтології в інтерфейсі об'єктного представлення.



Рейтинги установ - 2019															
Поисковый запрос	Загальний рейтинг установ	Вид роботи	Баллы	Количество (доля %)	Аудиторы	Риск	Часы учебного времени на бюджетно	Часы учебного времени на бюджетно	Общая цена	Результат	Всего за сотрудничество	Коэффициент вливания за написание	Загальний рейтинг виплатки	Загальний рейтинг науково-педагогічного працівника	Загальний рейтинг кафедри
Рейтинги установ	302,699	аспірантура, українське мовою	13	106						241,8					
Рейтинги установ	2019 Виплатка 1	Дисципліна забезпечення підготовки та проведення навчальних занять кося			аспірант, бакалавр, магістр	українське мовою	366,6	14,1	602,7	793,91					
Рейтинги установ	2019 Виплатка 2	бакалавр, українське мовою	13	206						371,8					
Рейтинги установ	2019 Виплатка 2	магістр, спеціальність, українське мовою	13	160						218,4					
Рейтинги установ	2019 Виплатка 2	програма навчальної дисципліни, нова дисципліна, Інтернет-система, розроблені системи, СКР-магістр, назва спеціальності: "Технології" та програма навчальної дисципліни, нова дисципліна, Системи розроблені, обробки інформації в телекомунікаційних мережах, СКР-магістр, програма навчальної дисципліни, нова дисципліна, Назва дисципліни, Назва дисципліни, Назва дисципліни за тематикою магістрської дисципліни: СКР-магістр, назва спеціальності	50	[100%]						50					
Рейтинги установ	2019 Виплатка 2	Розроблені програми навчальних дисциплін, програми кредитних модулів	50	[100%]						50					
Рейтинги установ	2019 Виплатка 2	2 Робота №2													
Рейтинги установ	2019 Виплатка 2	2 Робота №3													
Рейтинги установ	2019 Виплатка 2	2 Робота №4													
Рейтинги установ	2019 Виплатка 2	2 Робота №5													
Рейтинги установ	2019 Виплатка 2	2 Робота №6													
Рейтинги установ	2019 Виплатка 2	2 Робота №7													
Рейтинги установ	2019 Виплатка 2	2 Робота №8													
Рейтинги установ	2019 Виплатка 2	2 Робота №9													
Рейтинги установ	2019 Виплатка 2	2 Робота №10													
Рейтинги установ	2019 Виплатка 2	2 Робота №11													
Рейтинги установ	2019 Виплатка 2	2 Робота №12													
Рейтинги установ	2019 Виплатка 2	2 Робота №13													
Рейтинги установ	2019 Виплатка 2	2 Робота №14													
Рейтинги установ	2019 Виплатка 2	2 Робота №15													
Рейтинги установ	2019 Виплатка 2	2 Робота №16													
Рейтинги установ	2019 Виплатка 2	2 Робота №17													
Рейтинги установ	2019 Виплатка 2	2 Робота №18													
Рейтинги установ	2019 Виплатка 2	2 Робота №19													
Рейтинги установ	2019 Виплатка 2	2 Робота №20													
Рейтинги установ	2019 Виплатка 2	2 Робота №21													
Рейтинги установ	2019 Виплатка 2	2 Робота №22													
Рейтинги установ	2019 Виплатка 2	2 Робота №23													
Рейтинги установ	2019 Виплатка 2	2 Робота №24													
Рейтинги установ	2019 Виплатка 2	2 Робота №25													
Рейтинги установ	2019 Виплатка 2	2 Робота №26													
Рейтинги установ	2019 Виплатка 2	2 Робота №27													
Рейтинги установ	2019 Виплатка 2	2 Робота №28													
Рейтинги установ	2019 Виплатка 2	2 Робота №29													
Рейтинги установ	2019 Виплатка 2	2 Робота №30													
Рейтинги установ	2019 Виплатка 2	2 Робота №31													
Рейтинги установ	2019 Виплатка 2	2 Робота №3													

До складу онтологічної моделі входять рейтинги 16-ти викладачів за 2019 рік, та 2-х за 2018 та 2019.

61

Висновки

1. Розглянуто програмні середовища проектування онтологічних моделей
2. Розглянуто характерні особливості програмної платформи ТОДАОС
3. Описано фактичну розробку онтології представлення інформації наукових установ на базі ТОДАОС

ЗАГАЛЬНІ ВИСНОВКИ ПО РОБОТІ

1. Проведено огляд та порівняльний аналіз моделей представлення інформації та наведено класифікацію таких методів. Наведено недоліки та переваги кожного методу.
2. Проведено огляд процесу онтологічного моделювання інформації, визначено особливості використання цього методу представлення інформації.
3. Проведено аналіз показників діяльності наукових та освітніх установ, та критерії оцінки ефективності їх діяльності.
4. Запропоновано модель представлення різномірної слабоструктурованої інформації наукових та освітніх установ, що дозволить розробити онтологічну модель представлення інформації, яка дозволить підвищити ефективність оцінювання установ.
5. Описано елементи моделі представлення різномірної слабоструктурованої інформації за допомогою методу онтологічного моделювання.
6. Розроблено та описано онтологічну модель для представлення різномірної слабоструктурованої інформації наукових та освітніх установ засобами платформи ТОДАОС.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Глоба Л.С., Задоевко Б.О. ОЦІНЮВАННЯ ЕФЕКТИВНОСТІ ДІЯЛЬНОСТІ НАУКОВИХ УСТАНОВ МЕТОДОМ ОБРОБКИ РІЗНОРІДНОЇ СЛАБОСТРУКТУРОВАНОЇ ІНФОРМАЦІЇ Режим доступу: <http://conferenc.its.kpi.ua/2020/paper/view/20683/10878>
2. Змитрович А.И. Интеллектуальные информационные системы / А.И. Змитрович. – Мн.: НТООО «ТетраСистемс», 1997. – 368 с.
3. Круглов В.В. Интеллектуальные информационные системы: компьютерная поддержка систем нечеткой логики и нечеткого вывода / В.В. Круглов, М.И. Дли. – М.: Издательство Физико-математической литературы, 2002. – 256 с., 12. 12. Пупков К.А. Интеллектуальные системы / К.А. Пупков, В.Г. Коньков. – М.: Издательство МГТУ им. Н.Э. Баумана, 2003. – 348 с.
4. Helbig H. Knowledge Representation and the Semantics of Natural Language / H. Helbig. – Berlin: Springer, 2006. – 650 p.
5. Джордж Ф. Люгер Искусственный интеллект: стратегии и методы решения сложных проблем, 4-е издание: Пер. с англ. / Джордж Ф. Люгер. – М.: Издательский дом "Вильямс", 2003. – 864 с.
6. Макаллистер Дж. Искусственный интеллект и Пролог на микроЭВМ / Пер. с англ. А. В. Чукашова, М. В. Сергиевского; Под ред. М. В. Сергиевского. – М.: Машиностроение, 1990. – 240 с.
7. Мариничева М.К.: Управление знаниями на 100%: Путеводитель для практиков / М.К. Мариничева. – М.: Альпина Бизнес Букс, 2008. – 320 с.
8. Онтология_(информатика) [Электронный ресурс]. Энциклопедия Википедия. – Режим доступа : [http://ru.wikipedia.org/wiki/Онтология_\(информатика\)](http://ru.wikipedia.org/wiki/Онтология_(информатика))
9. Gruber T.R. A translation approach to portable ontologies [Электронный ресурс] / T.R. Gruber // Knowledge Acquisition. 1993. – № 5(2). – С. 199-

220. – Режим доступа : <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>
10. Gruber T.R. Ontology to appear in the Encyclopedia of Database Systems [Электронный ресурс] / Ling Liu and M. Tamer Özsu (Eds.). – Springer-Verlag, 2008. – Режим доступа : <http://tomgruber.org/writing/ontologydefinition-2007.htm>
11. Онтология [Электронный ресурс]. Энциклопедия Википедия. – Режим доступа : <http://ru.science.wikia.com/wiki/Онтология>
12. Онтологии и тезаурусы : учебное пособие / [Соловьев В.Д., Добров Б.В., Иванов В.В., Лукашевич Н.В.]. – Казань ; Москва, 2006. – 173 с.
13. Джарратано Дж. Экспертные системы: принципы разработки и программирование / Дж. Джарратано, Г. Райли. – М. : Вильямс, 2007. – 1152 с.
14. Ю.К. Шакирова Ю.К. Системы поддержки принятия решений в образовании / Ю.К. Шакирова, А.Е. Маденова, Н.К. Савченко, Г.Б. Абилдаева, Е.А. Воробьева // Дистанционное и виртуальное обучение. – 2016. – №2. – С. 69-73
15. Guarino N. Formal Ontology in Information Systems / N. Guarino // Proceedings of FOIS'98. – 1998. – Amsterdam: IOS Press. – P. 3-15
16. Gruber T. Toward Principles for the Design of Ontologies Used for Knowledge Sharing / T. Gruber // International Journal of Human-Computer Studies. – 1995. – Vol. 43. Issues 5–6. – P. 907–928
17. Guariano N. Ontologies and Knowledge Bases. Towards a Terminological Clarification / N. Guariano, P. Giaretta // Towards Very Large Knowledge Bases. – 1995. – IOS Press, Amsterdam. – P. 25–32
18. Guariano N. The Ontological Level: Revisiting 30 Years of Knowledge Representation / N. Guariano // Conceptual Modeling: Foundations and Applications. – 2009. – Berlin, Heidelberg: Springer-Verlag. – P. 52 – 67
19. Боровикова О.И. Подход к автоматизации сбора онтологической информации для интернет-портала знаний / О.И. Боровикова, Ю.А.

- Загорулько, Е.А. Сидорова // Компьютерная лингвистика и интеллектуальные технологии: Труды международной конференции Диалог'2005. – 2005. – Москва: Наука. – С. 65–70].
20. Globa L. S., Novogradskaya R. L. and Koval A. V. (2018) Ontology Model of Telecom Operator Big Data. Proceedings of IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), 2018. IEEE Digital Library, 8433710
21. Domingue J. Tadzebaev and WebOnto: Discussing, Browsing, and Editing Ontologies on the Web // Proc. of the Eleventh Workshop on Knowledge Acquisition, Modeling and Management, KAW'98, Banff, Canada — 1998.
22. Musen, M. Domain Ontologies in Software Engineering: Use of Protégé with the EON Architecture // Methods of Information in Medicine. – 1998. – P.540-550.
23. MacGregor R. Inside the LOOM classifier // SIGART bulletin, — 1991. — Vol.3, No.2, — P.70-76.
24. Стрижак О.Є. Трансдисциплінарна інтеграція інформаційних ресурсів [Текст]: автореф. дис. ... д-ра техн. наук: 05.13.06 / Стрижак Олександр Євгенійович; Нац. акад. наук України, Ін-т телекомунікацій і глобал. інформ. простору. — К., 2014. — 47 с.
25. Transdisciplinarity: Basarab Nicolescu. Talks with Russ Volckmann, Integral Review Journal, 4, 2007. — P. 76.
26. Технологии анализа данных. Data Mining, Visual Mining, Text Mining, OLAP (2-е издание) / А.А. Барсегян, М.С. Куприянов, В.В. Степаненко, И.И. Холод. — 2 изд., перераб. и доп. — СПб.: БХВ-Петербург, 2007. — 384 с.
27. Klein M. and Bernstein A., “Searching for Services on the Semantic Web using Process Ontologies”, in The First Semantic Web Working Symposium (SWWS-1), Stanford, CA USA, 2001.